Lecture Notes
On
RADAR & T.V ENGINEERING

Prof. Prasanta Kumar Nayak
Associate Professor - ECE

Department of Electronics and Communications Engineering

SPINTRONIC TECHNOLOGY & ADVANCE RESEARCH

Affiliated to BPUT Odisha

7th Semester	REC7D006	Radar and TV Engineering	L-T-P 3-0-0	3 Credits

Module I

Radar : The Radar equation-Pulse Radar-CW Radar-CW Radar with non zero IF, equation for Doppler frequency- FM-CW Radar using sideband superhetrodyne receiver, MTI Radar-Delay line canceller, MTI Radar with power amplifier & power oscillator, Non coherent MTI Radar, Pulse Doppler Radar, Radar Transmitters. Radar Modulator-Block diagram. Radar receivers- noise figure, low noise front ends, Mixers – Different types of Displays – Duplexers- Branch type and balanced type. Navigation- Loop Antenna, Radio compass. Hyperbolic Systems of Navigation, LORAN – A. Distance Measuring Equipment . Instrument Landing System – Localizer, Glide Slope, Marker beacons.

Module II

Television: Scanning, Blanking and synchronisation, Picture signal - composite video signalVestigial sideband transmission-Principle of CCD Camera - Monochrome picture tube-Monochrome TV receivers- RF tuner ,VHF tuner- Video amplifier, IF section, Vestigial sideband correction- Video detectors, Sound signal separation, AGC, sync separation, horizontal and vertical deflection circuits, EHT generation. Colour TV system: Principle of colour signal transmission and reception, PAL, NTSC, SECAM (block schematic description), Picture tube – delta gun.

Module III

Digital TV: Digitized Video, Source coding of Digitized Video – Compression of Frames – DCT based – (JPED), Compression of Moving Pictures (MPEG). Basic blocks of MPEG2 and MPE4. Digital Video Broadcasting (DVB) – Modulation: QAM – (DVB-S, DVB-C), OFDM for Terrestrial Digital TV (DVB –T). Reception of Digital TV Signals (Cable, Satellite and terrestrial). Digital TV over IP, Digital terrestrial TV for mobile. Display Technologies – basic working of Plasma, LCD and LED Displays.

Books:

1. Merrill I. Skolnik: Introduction to Radar Systems,3/e, Tata McGraw Hill,
2. N.S.Nagaraja: Elements of Electronic Navigation, 2/e, Tata McGraw Hill
3. R.R. Gulati: Monochroeme and Colour Television. New Age international, 2008.
4. Herve Benoit, Digital Television Satellite, Cable, Terrestrial, IPTV, Mobile TV in the DVB Framework, 3/e, Focal Press, Elsevier, 2008
5. Shlomo Ovadia: Broadband Cable TV Access Networks, PH-PTR, 2001
6. Byron Edde: Radar Principles, Technology & Applications, Pearson Education.
7. Mark E Long: ―The Digital Satlitte TV Hand Book‖, Butterworth-Heinemann.
8. K.R.Rao, J.O.Hwang, Techniques and standards for Image,Video and Audio coding,Prentice Hall,1996
9. John Arnold, Michael Frater, Mark Pickering,Digital Television Technology and Standards, John Wiley & Sons, Inc, 2007
10. Robert L. Hartwig,Basic TV Technology: Digital and Analog, 4/e, Focal Press, Elsevier, 2005

INTRODUCTION TO RADAR SYSTEM

1.1 Introduction:-

Radar is an electromagnetic system for the detection and location of objects. It operates by transmitting a particular type of waveform, a pulse-modulated sine wave for example, and detects the nature of the echo signal. Radar is used to extend the capability of one's senses for observing the environment, especially the sense of vision.

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.

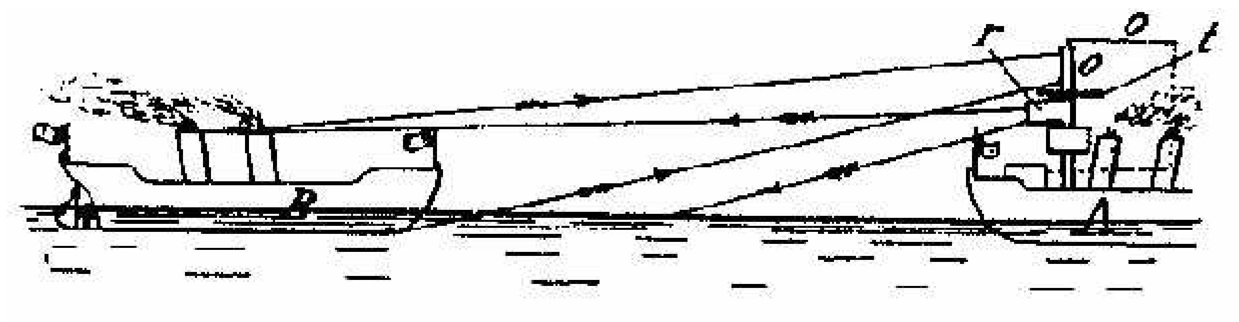
A portion of the transmitted signal is intercepted by a reflecting object (target) and is reradiated in all directions. It is the energy reradiated in the back direction that is of prime interest to the radar. The receiving antenna collects the returned energy and delivers it to a receiver, where it 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.

1.2 History Background

James Clerk Maxwell (1831 –1879) - predicted the existence of radio waves in his theory of electromagnetism. In 1886, Hertz experimentally tested the theories of Maxwell and demonstrated the similarity between radio and light waves. Hertz showed that radio waves could be reflected itself. Heinrich Hertz, in 1886, experimentally tested the theories of Maxwell and demonstrated the similarity between radio and light waves. Hertz showed that radio waves could be reflected by metallic and dielectric bodies. Due to these reflections occurred through metallic bodies given a start to the development of radar systems.

In 1903 a German engineer by the name of Hiilsmeyer experimented with the detection of radio waves reflected from ships. He obtained a patent in 1904 in several countries for an radio waves reflected from ships as shown in fig.1.



(a)

(b)

Fig. 1 (a) Detection of wooden ship in 1904 **(b)** Hülsmeyer 1904, who detected the first object through radar

In the autumn of 1922 A. H. Taylor and L. C. Young of the Naval Research Laboratory detected a wooden ship using a CW wave-interference radar with separated receiver and transmitter. The wavelength was 5 m. The first application of the pulse technique to the measurement of distance was in the basic scientific investigation by Breit and Tuve in 1925 for measuring the height of the ionosphere. However, more than a decade was to elapse before the detection of aircraft by pulse radar was demonstrated.

The first detection of aircraft using the wave-interference effect was made in June, 1930, by L. A. tlyland of the Naval Research Laboratory.' It was made accidentally while he was working with a direction-finding apparatus located in an aircraft on the ground. The transmitter at a frequency of 33 MHz was located 2 miles away, and the beam crossed an air lane from L. Hyland of the Naval Research Laboratory. It was made accidentally while he was working with a direction-finding apparatus located in an aircraft on the ground. The transmitter at a frequency of 33 MHz was located 2 miles away, and the beam crossed an air lane from a nearby airfield.

Before the advent of radar, the only practicable means of detection of aircraft was acoustic, and a network of acoustic detectors was built in the 1920s and 1930s around the south and east coast of the UK, some of which still remain. In calm air conditions, detection ranges of up to 25km were achievable.

(a)

(b) (c)

Fig. 2 Different types of Acoustic Radars from 1920-1930

Radar Applications:-

In aviation, aircraft are equipped with radar devices that warn of aircraft or other obstacles in or approaching their path, display weather information, and give accurate altitude readings. The first commercial device fitted to aircraft was a 1938 Bell Lab unit on some United Air Lines aircraft. Such aircraft can land in fog at airports equipped with radar-assisted ground-controlled approach systems in which the plane's flight is observed on radar screens while operators radio landing directions to the pilot.

Marine radars are used to measure the bearing and distance of ships to prevent collision with other ships, to navigate, and to fix their position at sea when within range of shore or other fixed references such as islands, buoys, and lightships. In port or in harbour, vessel traffic service radar systems are used to monitor and regulate ship movements in busy waters.

Normal radar functions:

1. Range (from pulse delay)
2. Velocity (from Doppler frequency shift)
3. Angular direction (from antenna pointing)

Signature analysis and inverse scattering:

4. Target size (from magnitude of return)
5. Target shape and components (return as a function of direction)
6. Moving parts (modulation of the return)
7. Material composition

The complexity (cost & size) of the radar increases with the extent of the functions that the radar performs.

BASIC PRINCIPLES OF RADAR

A radar system has a transmitter that emits radio waves called *radar signals* in moving or stationary target directions. When these come into contact with an object they are usually reflected or scattered in many directions. Radar signals are reflected especially well by materials of considerable electrical conductivity especially by most metals, by seawater and by wet ground. Some of these make the use of radar altimeters possible. The radar signals that are reflected back towards the transmitter are the desirable ones that make radar work. If the object is *moving* either toward or away from the transmitter, there is a slight equivalent change in the frequency of the radio waves, caused by the Doppler effect.

The basic principle of the radar is shown in fig. 2.1. A transmitter generates an electromagnetic signal that is radiated by the antenna into space. A portion of the transmitted electromagnetic energy is reflected back by the target towards the radar. Based on the received target echo signal the receiver made decision for the position, range and direction of the target. The term radar is a contraction of the words radio detection and ranging.

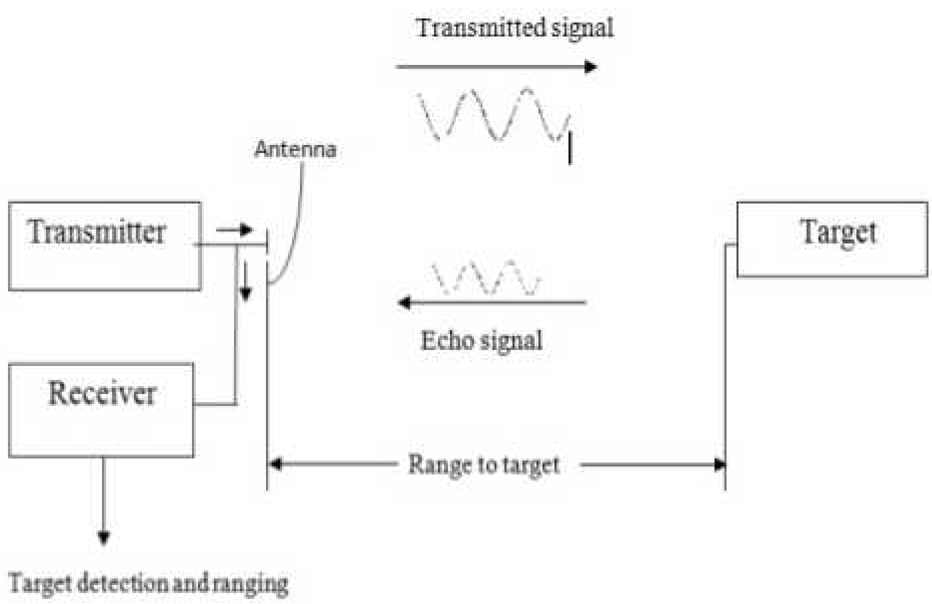


Fig. 2.1 Basic Principles of the Radar

The basic terminology used for radar is discussed as follows.

Range:- The range of the target is observed by measuring the time (T_R) it takes for the radar signal to travel to the target and return back to the radar. Thus the time for the signal to travel to the target located at range (R) and the return back to the radar is *2R/C*. The range of the target can be given as:

$$R = \frac{cT_R}{2} \qquad \dots (1)$$

with the range in kilometers or in nautical miles, and T in microseconds.

$$\begin{aligned} R(km) &= 0.15T_R\ (\text{«}s) \\ R(nmi) &= 0.081T_R\ (\text{«}s) \end{aligned} \qquad \dots (2)$$

Maximum Unambiguous Range:- Once a signal is radiated into space by a radar, enough time must elapse to allow all echo signal to return to the radar before the transmission of next pulse. The rate at which the pulses are transmitted, is determined by the longest range of the target. If the time between pulses T_P is too short, an echo signal from the long range target might arrive after the transmission of the next pulse. The echo that arrives after the transmission of next pulse is called as *second-time-around-echo (or multiple-time-around-echo*). Such an echo would appear to be at a closer range than actual, this range measurement will be misleading for range calculation, if it is not known that this is second time echo. The range beyond which the target appears as second-time-around-echoes is the *maximum unambiguous range*, R_{un} and is given by

$$\begin{aligned} R_{un} &= \frac{cT_P}{2} = \frac{c}{2f_p} \\ f_p &= \frac{1}{T_P} \qquad \dots (3) \\ Duty\,cycle &= \frac{2}{T_P} \end{aligned}$$

Where T_P is the pulse repletion time and f_p is the pulse repetition frequency.

A problem with pulsed radars and range measurement is how to unambiguously determine the range to the target if the target returns a strong echo. This problem arises because of the fact that pulsed radars typically transmit a sequence of pulses. The radar receiver measures the time between the leading edges of the last transmitting pulse and the echo pulse. It is possible that an echo will be received from a long range target after the transmission of a second transmitting pulse.

In this case, the radar will determine the wrong time interval and therefore the wrong range. The measurement process assumes that the pulse is associated with the second transmitted pulse and declares a much reduced range for the target. This is called range ambiguity and occurs where there are strong targets at a range in excess of the pulse repetition time. The pulse repetition time defines a maximum unambiguous range. To increase the value of the unambiguous range, it is necessary to increase the PRT, this means: to reduce the PRF.

Echo signals arriving after the reception time are placed either into the transmit time where they remain unconsidered since the radar equipment isn't ready to receive during this time, or into the following reception time where they lead to measuring failures (ambiguous returns).

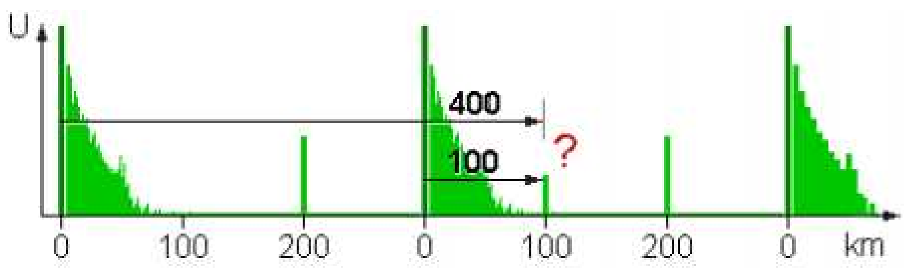


Fig. 2.2 a second-time-arounr-echo in a distance of 400 km assumes a wrong range of 100 km

Pulse Repetition Frequency (PRF):- The rate at which the pulses are transmitted towards the target from the radar is called as the pulse repletion frequency, f_p.

$$f_p = \frac{1}{T_P} \quad \text{… (4)}$$

Pulse Repetition Period:- The time interval at which the pulses are periodically transmitted towards the target from the radar is called as the pulse repletion period, T_P is given by in terms of prf.

$$T_p = \frac{1}{f_P} \quad \text{… (5)}$$

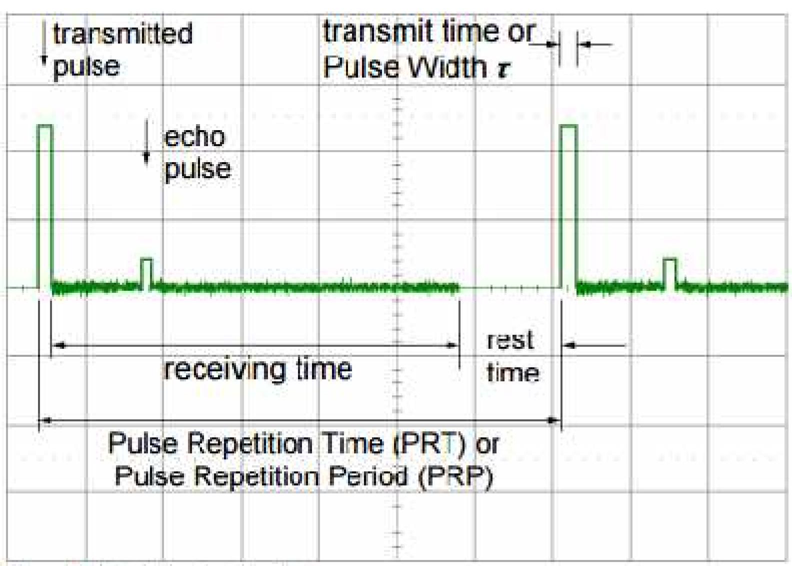


Fig. 2.3 A typical radar time line

Duty Cycle:- The duty cycle of the radar waveform is described as the ratio of the total time the radar is radiating to the total time it could have radiated.

$$Duty\,cycle = \frac{p_{av}}{P_T} \qquad \ldots (6)$$

$$Duty\,cycle = \frac{\tau}{T_P} = \tau\, f_p \qquad \ldots (7)$$

Where τ is pulse width of the transmitted pulse and T_P is the pulse repetition period.

Peak Power of the Radar:- The maximum power of the radar antenna, that can be transmitted for the maximum unambiguous range target detection in particular direction.

Average Power of the Radar:- The average power of the radar antenna, that can be transmitted for the maximum unambiguous range target detection in all the direction (for isotropic antenna).

Radar Wave forms:- Typical radar utilizes various waveforms for target detection.

- **Pulse waveform:-** A radar uses rectangular pulse wave form with pulse width of 1microsecond, pulse repletion period 1 millisecond.
- **Continuous waveform:-** A very long continuous waveform are required for some long range radars to achieve sufficient energy for small target detection.

o

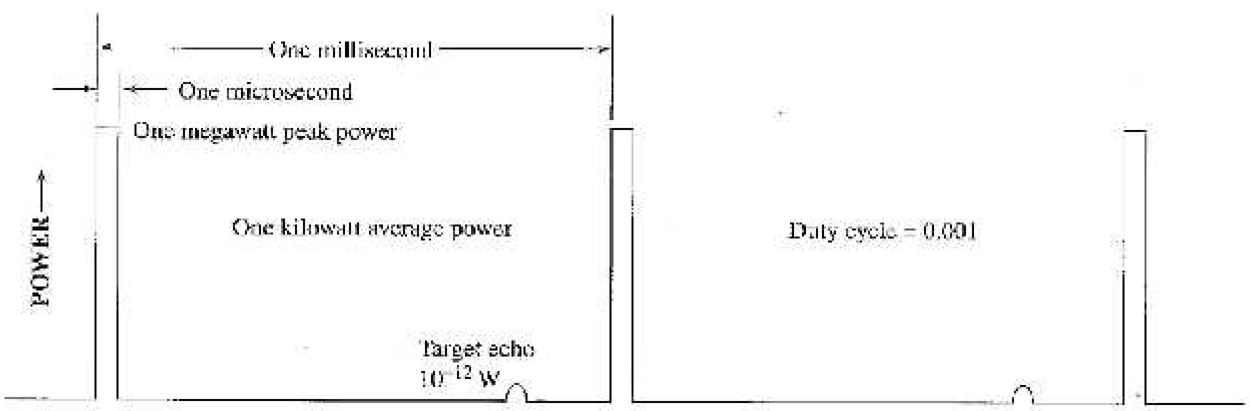


Fig.2.4 Example of typical pulse waveform for medium range air surveillance radar

RADAR RANGE EQUATION

3.1 Introduction:-

The radar range relates the radar range with the characteristics of transmitter, receiver antenna, target and environment. The radar range equation is useful to understand the maximum range of the radar that can be detected by the radar with their performance parameters. One of the simpler equations of radar theory is the radar range equation.

3.2 BASIC RADAR RANGE EQUATIONS

The transmitted power P_t is radiated by an isotropic antenna, the power density at distance R can be given as:

$$\text{Power demnsity at range R from an isotropic antenna} = \frac{P_t}{4\text{®} R^2} \text{ (Watt/square meter)} \quad \ldots (3.1)$$

The maximum gain of the antenna can be defined as:

$$G = \frac{\text{max power density radiated byan antenna}}{\text{power density radiated by a lossless isotropic antenna}} \quad \ldots (3.2)$$

Thus the power density at target from a directive antenna can be given as:

$$\text{Power density at range R from } a \text{ directive antenna} = \frac{P_t G}{4\text{®} R^2} \quad \ldots (3.3)$$

The target receives a portion of the incident energy and reflected it in various directions. Thus the radar cross section of the target determines the power density returned back to the radar.
The reflected power from the target through its cross section (target cross section) can be given as:

$$\text{Reflected power from the target towards the radar} = \frac{P_t G}{4\text{®} R^2} \bullet \frac{\pm}{4\text{®} R^2} \quad \ldots (3.4)$$

The radar antenna receives a portion of the reflected power from the target cross section. the received power can be given as:

$$P_r = \frac{P_t G}{4\text{®} R^2} \bullet \frac{\pm}{4\text{®} R^2} \bullet A_e \quad \ldots (3.5)$$

$$A_e = \eta_a \cdot A \qquad \ldots (3.6)$$

Where A_e is the effective area of the receiving antenna, A is the physical antenna area and η_a is the antenna aperture efficiency. The maximum range of the radar (Rmax) can be defined as the maximum distance beyond which radar cannot detect the target. So the received signal power can be given as the minimum detectable signal.

$$S_{\min} = \frac{P_t G}{4\pi R^2} \cdot \frac{\sigma}{4\pi R_{\max}{}^2} \cdot A_e \qquad \ldots (3.7)$$

$$R_{\max} = \left[\frac{P_t G}{4\pi} \cdot \frac{\sigma}{4\pi} \cdot \frac{A_e}{S_{\min}} \right]^{1/4} \qquad \ldots (3.8)$$

This is the fundamental form of radar range equation. If the antenna is used for both the transmission and receiving purpose, then the transmitted gain (G) can be given in terms of the effective area (A_e).

$$G = \frac{4\pi A_e}{\lambda^2} \qquad \ldots (3.9)$$

Now the maximum radar range can be given as follows.

$$R_{\max} = \left[\frac{P_t G^2 \lambda}{(4\pi)^3} \cdot \sigma \cdot \frac{A_e}{S_{\min}} \right]^{1/4} \quad \text{(When G is constant)} \qquad \ldots (3.10)$$

$$R_{\max} = \left[\frac{P_t}{(4\pi)^3} \cdot \sigma \cdot \frac{A_e{}^2}{S_{\min}} \right]^{1/4} \quad \text{(When } A_e \text{ is constant)} \qquad \ldots (3.11)$$

These three forms of radar range equations [2.8, 2.10 and 2.11] are based on the effective area (A_e) and transmitter antenna gain (G).

3.3 Radar Block Diagram

The operation of a typical pulse radar may be described with the aid of the block diagram shown in Fig. 1.2. The transmitter may be an oscillator. such as a magnetron. that is "pulsed" (turned on and off) by the modulator to generate a repetitive train of pulses. The magnetron has prohahly heen the most widely used of the various microwave generators for radar. A typical radar for the detection of aircraft at ranges of 100 or 200 nmi might employ a peak power of the order of a megawatt. an average power of several kilowatts, a pulse width of several microseconds. and a

pulse repetition frequency of several hundred pulses per second. 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 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.

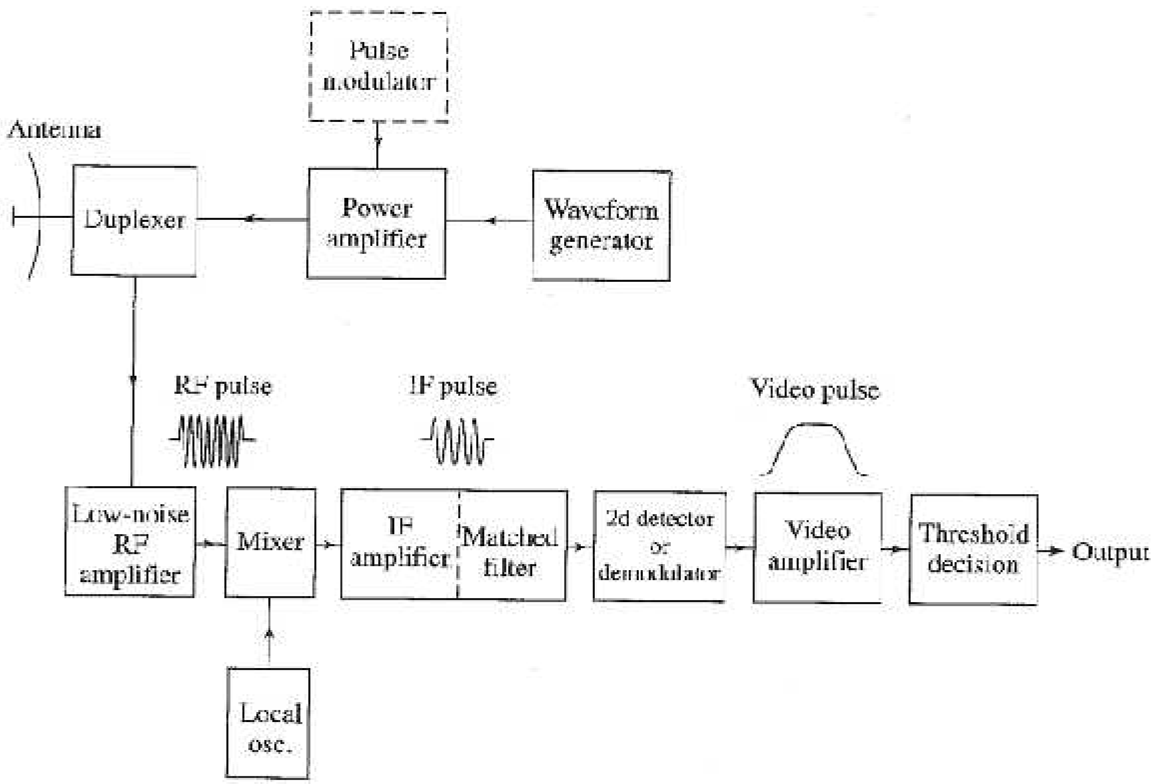


Fig. 3.1 Radar Block Diagram

The mixer and local oscillator (LO) convert the RF signal to an intermediate frequency (IF). A "typical" IF amplifier for an air-surveillance radar might have a center frequency of 30 or 60 MHz and a bandwidth of the order of one megahertz.

The IF amplifier should be designed as a matchted filter; i.e., its frequency-response function H(f) should maximize the peak-signal-to-mean-noise-power ratio at the output.

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). Timing signals are also supplied to the indicator to provide the range zero. Angle information is obtained from the pointing direction of the antenna.

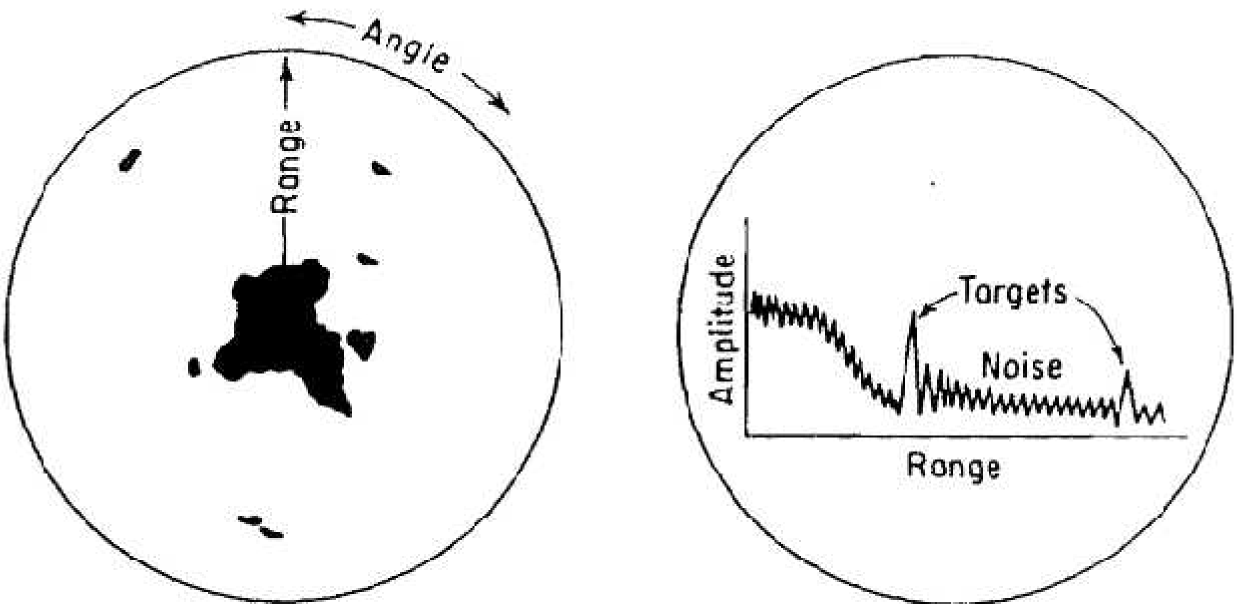


Fig. 3.2 (a) PPI presentation displaying range vs. angle (intensity modulation); (b) A scope presentation displaying amplitude vs. range (deflection modulation).

A common form of radar antenna is a reflector with a parabolic shape, fed (illuminated) from a point source at its focus. The parabolic reflector focuses the energy into a narrow beam, just as does a searchlight or an automobile headlamp. The beam may be scanned in space by mechanical pointing of the antenna. Phased-array antennas have also been used for radar. In a phased array the beam is scanned by electronically varying the phase of the currents across the aperture.

3.3 Radar's Electromagnetic Spectrum

Conventional radars generally have been operated at frequencies extending from about 220 MHz to 35 GHz, a spread of more than seven octaves. These are not necessarily the limits, since radars can be, and have been, operated at frequencies outside either end of this range.Skywave HF over-the-horizon (OTH) radar might be at frequencies as low as 4 or 5 MHz, and Groundwave HF radars as low as 2 MHz. At the other end of the spectrum, millimeter radars Have operated at 94 GHz. Laser radars operate at even higher frequencies.

The place of radar frequencies in the electromagnetic spectrum is shown in Fig. 3.3. Some of the nomenclature employed to designate the various frequency regions is also shown. Early in the development of radar, a letter code such as S, X, L, etc., was employed to Designate radar frequency bands. Although its original purpose was to guard military secrecy, the designations were maintained, probably out of habit as well as the need for some convenient short nomenclature. This usage has continued and is now an accepted practice of radar engineers.

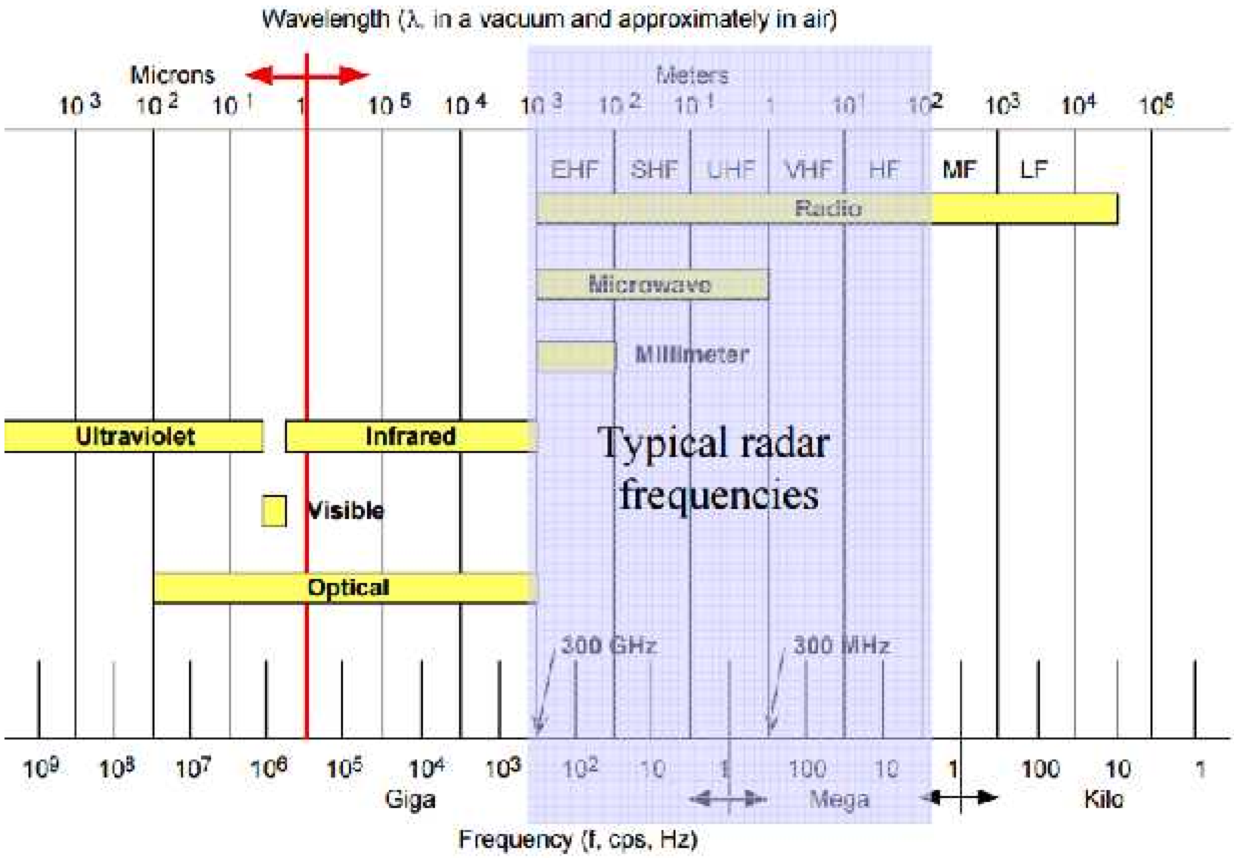


Fig. 3.3 Frequency spectrum for radar frequencies

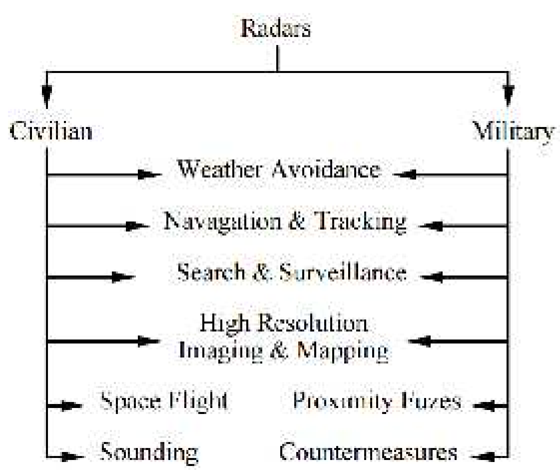
Table 3.1 lists the radar-frequency letter-band nomenclature adopted by the IEEE. These are related to the specific bands assigned by the International Telecommunications Union for radar. For example, although the nominal frequency range for L band is 1000 to 2000 MHz, an L-band radar is thought of as being confined within the region from 1215 to 1400 MHz since that is the extent of the assigned band.

Table 3.1 Radar Bands and their Usage

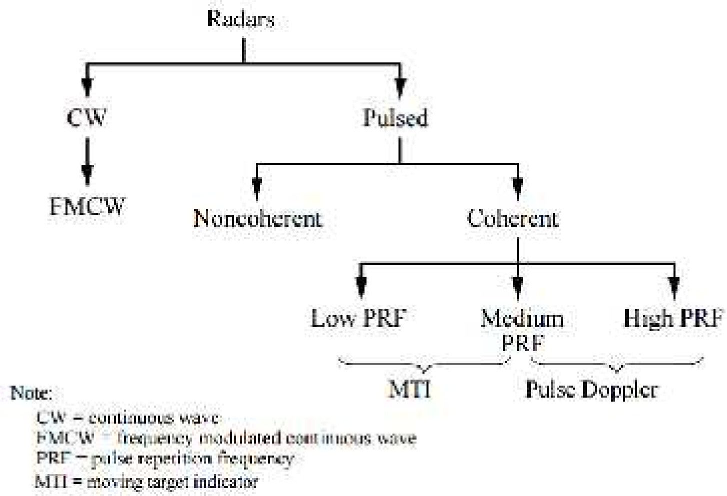
Band Designation	Frequency Range	Usage
HF	3–30 MHz	OTH surveillance
VHF	30–300 MHz	Very-long-range surveillance
UHF	300–1,000 MHz	Very-long-range surveillance
L	1–2 GHz	Long-range surveillance En route traffic control
S	2–4 GHz	Moderate-range surveillance Terminal traffic control Long-range weather
C	4–8 GHz	Long-range tracking Airborne weather detection
X	8–12 GHz	Short-range tracking Missile guidance Mapping, marine radar Airborne intercept
K_u	12–18 GHz	High-resolution mapping Satellite altimetry
K	18–27 GHz	Little use (water vapor)
K_a	27–40 GHz	Very-high-resolution mapping Airport surveillance
millimeter	40–100+ GHz	Experimental

3.4 Radar classification

Radar can be classified based on the function and the waveforms



(a)



(b)

Fig. 3.3 Radar can be classified based on the (a) function and (b) waveforms

In practice, however, the simple radar equation does not predict the range performance of actual radar equipments to a satisfactory degree of accuracy. The predicted values of radar range are usually optimistic. In some cases the actual range might be only half that predicted. Part of this discrepancy is due to the failure of Eq. (3.10) to explicitly include the various losses that can occur throughout the system or the loss in performance usually experienced when electronic equipment is operated in the field rather than under laboratory-type conditions &.another important factor that must be considered in the radar equation is the statistical or unpredictable nature of several of the parameters. The minimum detectable signal Smin and the target cross section (σ) are both statistical in nature and must be expressed in statistical terms.

3.5 MINIMUM DETECTABLE SIGNAL

The ability of a radar receiver to detect a weak echo signal is limited by the noise energy that occupies the same portion of the frequency spectrum as does' the signal energy. The weakest signal the receiver can detect is called. the minimum detectable signal. The specification of the minimum detectable signal is sometimes difficult because of its statistical nature and because the criterion for deciding whether a target is present or not may not be too well defined.

Detection is based on establishing a threshold level at the output of the receiver. If the Receiver output exceeds the threshold, a signal is assumed to be present. This is called threshold detection.

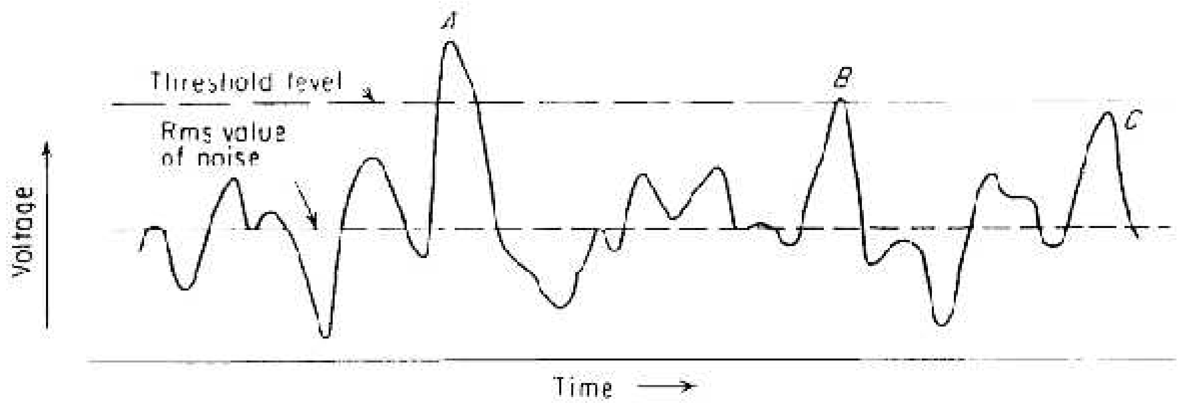


Fig. 3.4 Typical envelope of the radar receiver output as a function of time. A, and B, and C represent signal plus noise. A and B would be valid detections, but C is a missed detection.

A target is said to be detected if the envelope crosses the threshold. if the signal is large such as at A, it is not difficult to decide that a target is present. But consider the two signals at B and C, representing target echoes of equal amplitude. The noise voltage accompanying the signal at B is large enough so that the combination of signal plus noise exceeds the threshold.

Weak signals such as C would not be lost if the threshold level were lower. But too low a threshold increases the likelihood that noise alone will rise above the threshold and be taken for a real signal. Such an occurrence is called a false alarm.

CONTINUOUS WAVE AND FREQUENCY MODULATED RADAR

4.1 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/λ. 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 $4\pi R/\lambda$.

If target is in motion the range R and phase φ is continually changing. Thus the change in phase with respect to time can be given as frequency.

$$\frac{d\phi}{dt} = \frac{4\pi}{\lambda}\frac{dR}{dt} \qquad \ldots (1)$$

Range with respect to time can be defined as the radial velocity of the target. Thus the Doppler angular frequency can be given as:

$$\omega_d = 2\pi f_d = \frac{4\pi}{\lambda} v_r \qquad \ldots (2)$$

Where f_d is Doppler frequency and v_r is the radial velocity of the target with respect to radar. The Doppler frequency can be related with transmitter frequency f_0.

$$f_d = \frac{2v_r}{\lambda} = \frac{2v_r f_0}{c} \qquad \ldots (3)$$

When v_r is given in knots then the Doppler frequency can be given as:

$$f_d = \frac{1.03 v_r\,(knots)}{\lambda\,(m)} \qquad \ldots (4)$$

The relative velocity may be written $v_r = v\cos\theta$ where v is the target speed and θis the Angle made by the target trajectory and the line joining radar and target. When θ = 0, the doppler frequency is maximum. The doppler is zero when the trajectory is perpendicular to the radar line of sight (θ= 90°).

A plot of doppler frequency shifts as a function of radial velocity and the radar frequency bands is given in fig. 4.2. This figure illustrates that as the target radial velocity get increases the Doppler frequency shifts get increases with higher radar frequencies.

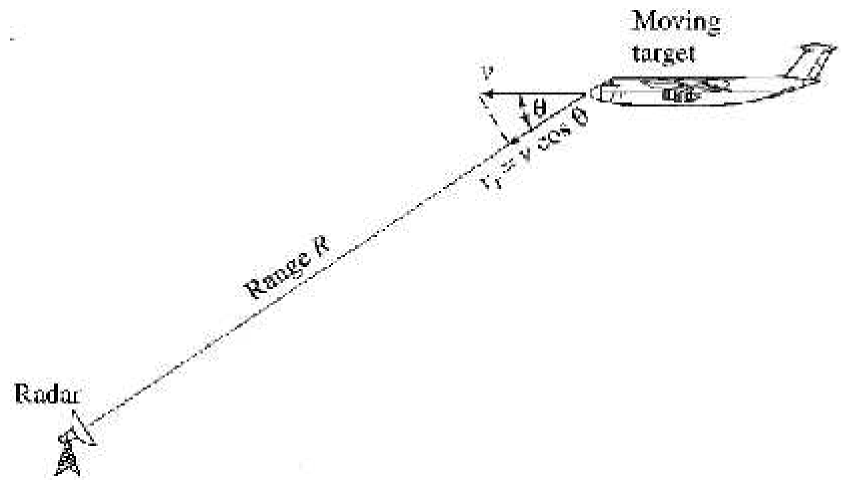


Fig. 4.1 Geometry of Radar and target in deriving the Doppler shifts

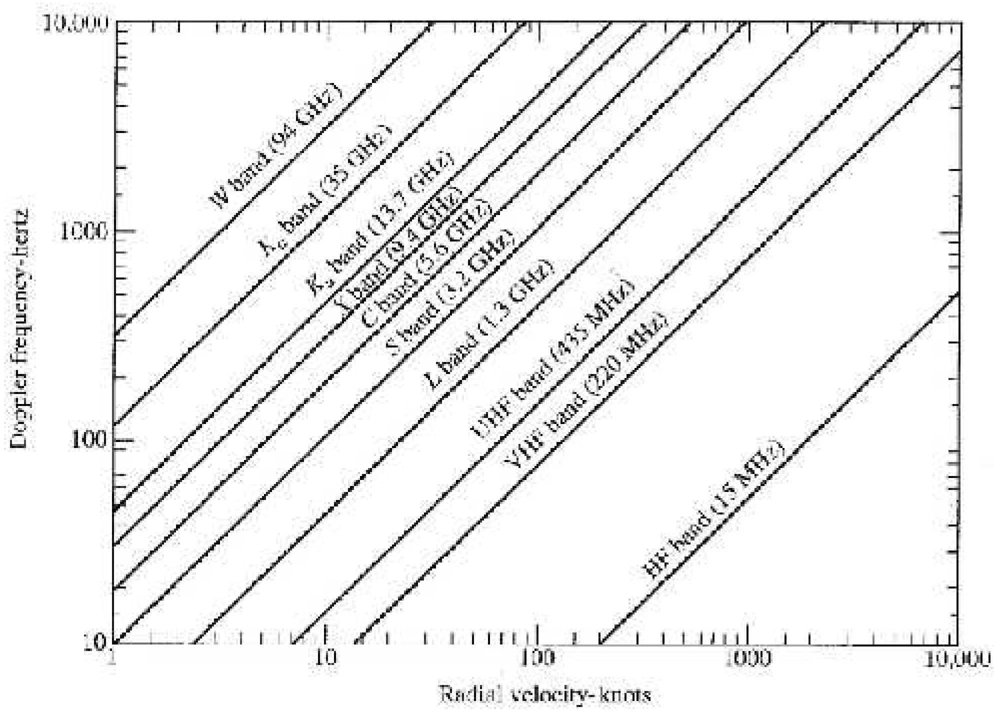


Fig. 4.2 Doppler frequency shifts for a moving target as a function of v_r and radar frequency band.

4.2 Continuous Wave Radar (CW Radar):-

A block diagram of simple CW radar is shown in Fig. 4.3. The transmitter generates a continuous (unmodulated) 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 $\pm f_d$ as given by Eq. (4).

- 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/o to produce a doppler beat note of frequency f_d The sign f_d is lost in this process.

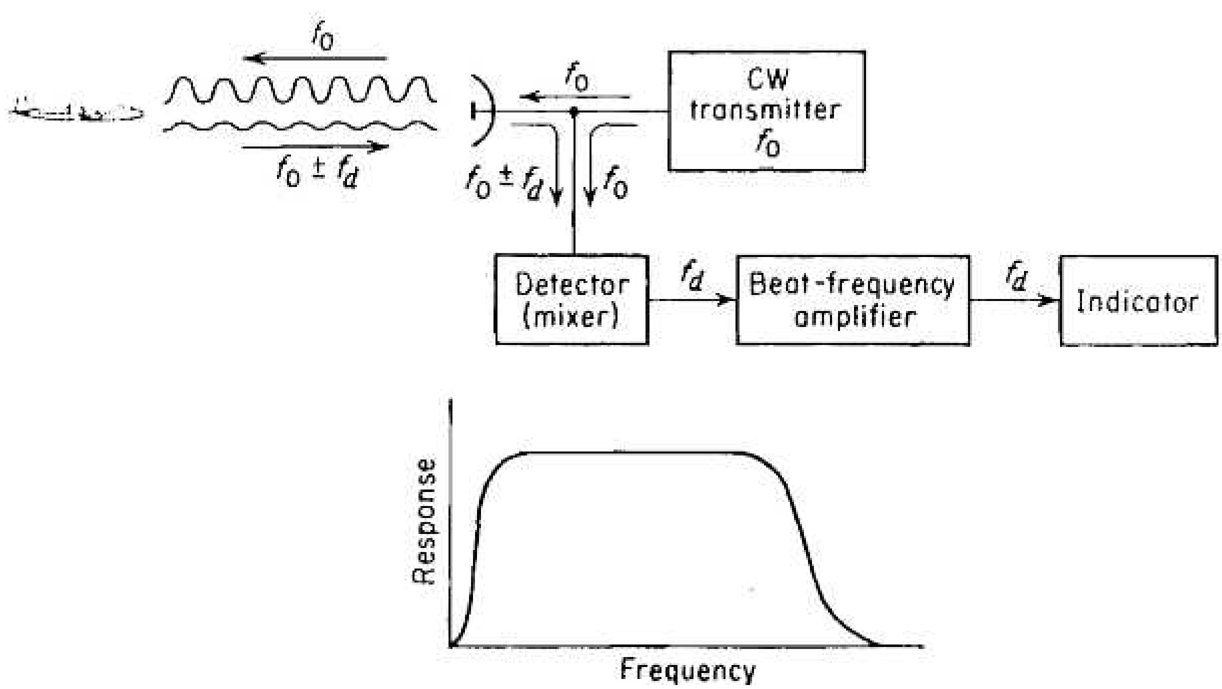


Fig. 4.3 CW Radar with frequency response

Pulse Radar: Pulse radar that extracts the Doppler frequency-shifted echo signal. A simple way to convert the CW radar to the pulse radar by turning on and off CW oscillator to generate pulses. This way of generation of pulses removes the reference signal, which is required to recognize the Doppler shifts. One way to introduce the reference signal is shown in fig. 4.4. Here the power amplifier is turned on and off to generate the high power pulses. The received echo signal is mixed with the output of CW oscillator, which acts as coherent reference to allow the recognition of any change in the frequency. Here coherent means that the transmitted pulses are synchronously used as reference signal. The change in frequency is detected through Doppler filter.

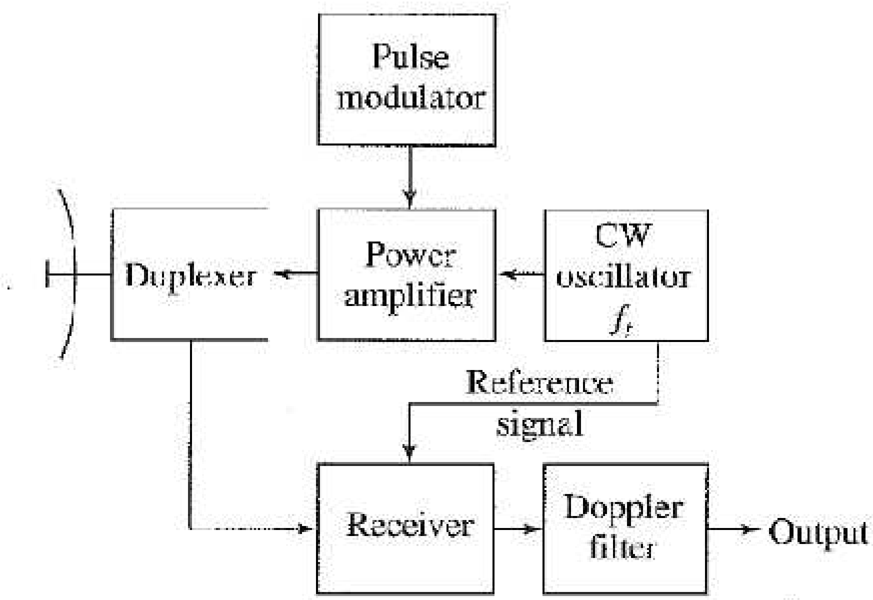


Fig. 4.4 Block diagram of simple Pulse Radar

Sweep to sweep subtraction:

The bipolar video (signal has positive and negative values) from two successive sweeps of MTI radar is shown in fig. 4.5. If one sweep is subtracted from the previous sweep, fixed clutter echoes will get cancel, and will not detected. On the other hand, moving target change its amplitude from sweep to sweep due the Doppler frequency shift. If one sweep is subtracted from other, the result will be canceled residue as shown in fig. 3.5.

Subtraction of the echoes from two successive sweeps is accomplished in delay line cancellers as shown in fig. 4.6. The delay-line canceller acts as a filter to eliminate the doc component of fixed targets and to pass the a-c components of moving targets. The video portion of the receiver is divided into two channels. One is a normal video channel. In the other, the video signal experiences a time delay equal to one pulse-repetition period (equal to the reciprocal of the pulse-repetition frequency). The outputs from the two channels are subtracted from one another. The fixed targets with unchanging amplitudes from pulse to pulse are canceled on subtraction.

However, the amplitudes of the moving-target echoes arc not constant from pulse 10 pulse, and subtraction results in an uncanceled residue.

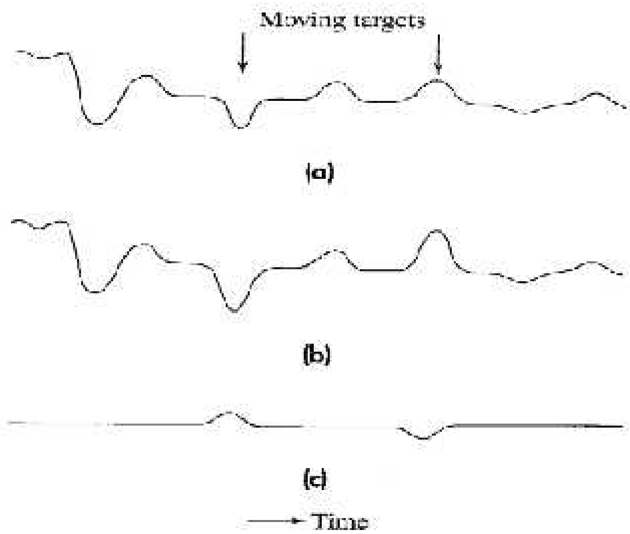


Fig. 4.5 Sweep to sweep subtraction

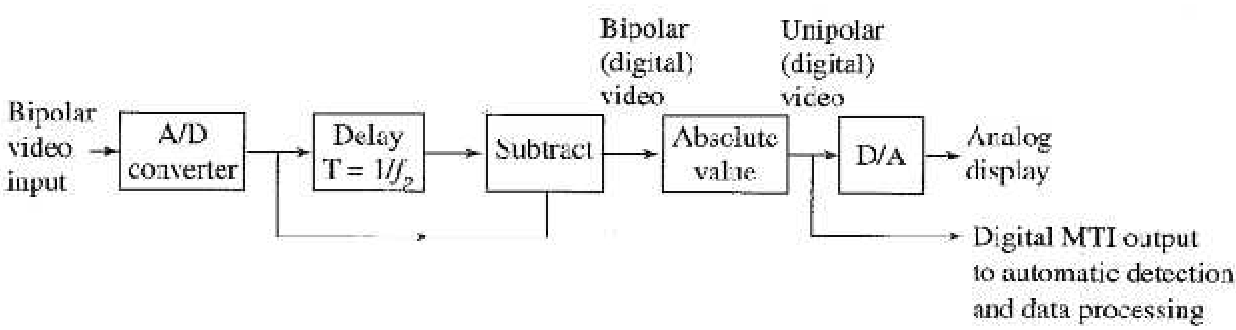


Fig. 4.6 Block diagram of single delay line canceller

MTI Radar Block Diagram:-

The doppler frequency shift [Eq. (3.2)] produced by a moving target may be used in a pulse radar. just as in the CW radar discussed in Chap. 3, to determine the relative velocity of a target or to separate desired moving targets from undesired stationary objects (clutter). 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 block diagram of a more common MTI radar employing a power amplifier is shown in Fig. 4.5. The significant difference between this MTI configuration is the manner in which the reference signal is generated. In Fig. 4.7, the coherent reference is supplied by an oscillator called the coho, which stands for coherent oscillator.

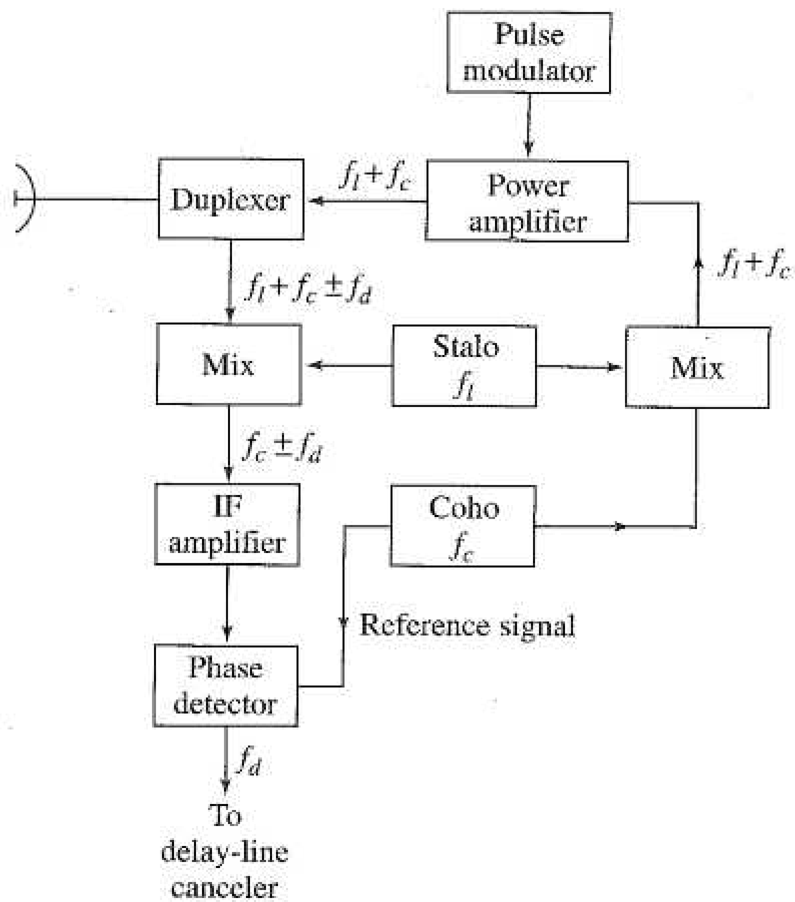


Fig. 4.7 Block diagram of MTI radar with power-amplifier transmitter

- The coho is a stable oscillator whose frequency is the same as the intermediate frequency used in the receiver. In addition to providing the refcrence signal. the output of the coho. fc is also mixed with the local-oscillator frequency fl.

- The local oscillator must also be a stable oscillator and is called stalo, for stable local oscillator.

- The stalo, coho, and the mixer in which they are combined plus any low-level amplification are called the receiver-exciter because of the dual role they serve in both the receiver and the transmitter.

- 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 reference signal from the coho and the I F 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.

Delay Line Canceller:-

The simple MTI delay-line canceller shown in Fig. 4.6 is an example of a time-domain filter.The capability of this device depends on the quality of the medium used as the delay line. Thedelay 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 delay line must introduce a time delay equal.

One of the advantages of a time-domain delay-line canceler 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.

Frequency Response of Delay Line canceller

The delay-line canceler acts as a filter which rejects the d-c 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 signal from a target at range R_0, the output of the phase detector can be given as:

$$V_1 = k\sin(2\pi f_d t - \phi_0) \qquad \ldots (5)$$

Where f_d is Doppler frequency, ϕ_0 constant phase of $4\pi R_0/\lambda$. The signal from the previous radar transmission is similar, which is delayed by time T_P

$$V_2 = k\sin[2\pi f_d(t - T_P) - \phi_0] \qquad \ldots (6)$$

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

$$V = V_1 - V_2 = 2k\sin(\pi f_d T_P)\cos\left[2\pi f_d\left(t - \frac{T_P}{2}\right) - \phi_0\right] \qquad \ldots. (7)$$

The magnitude of the relative frequency-response of the delay-line canceler [ratio of the amplitude of the output from the delay-line canceler, $2k\sin(\pi f_d T_P)$, to the amplitude of the normal radar video k] is shown in Fig. 4.8.

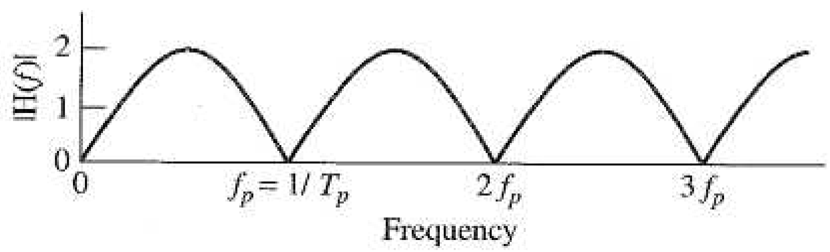


Fig. 4.8 Frequency response of the single delay-line canceller; T = delay time $=1/f_p$

Blind Speed:-

The response of the single-delay-line canceler will be zero whenever the argument $\pi f_d T_P$ in the amplitude factor of Eq. (7) is 0, π, 2 π, ... , etc., or when

$$f_d = \frac{2V_r}{\lambda} = \frac{n}{T_P} = nf_p \quad n = 0,1,2,3,\ldots\ldots \qquad \ldots (8)$$

The delay-line cancela 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

there of. Those relative target velocities which result in zero MTI response are called blind speed and can be given as:

$$v_n = \frac{n\lambda}{2T_p} = \frac{n\lambda f_p}{2} \quad n = 0,1,2,3,\ldots\ldots \qquad \ldots (9)$$

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

$$v_n = \frac{n^{a} f_p}{1.02} \approx n^{a} f_p \qquad \ldots (10)$$

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.

Pulse Doppler Radar:-

A pulse radar that extracts the doppler frequency shift for the purpose of detecting moving targets in the presence of clutter is either an MTI radar or a pulse doppler radar.

The distinction between them is based on the fact that in a sampled measurement system like a pulse radar, ambiguities can arise in both the doppler frequency (relative velocity) and the range (time delay) measurements. Range ambiguities are avoided with a low sampling rate (low pulse repetition frequency), and doppler frequency ambiguities are avoided with a high sampling rate. However, in most radar applications the sampling rate, or pulse repetition frequency, cannot be selected to avoid both types of measurement ambiguities.

The pulse doppler radar is more likely to use range-gated doppler filter-banks than delay-line cancelers. Also, a power amplifier such as a klystron is more likely to be used than a delay-line cancelers. A pulse doppler radar operates at a higher duty cycle than does an MTI. Although it is difficult to generalize, the MTI radar seems to be the more widely used of the two, but pulse doppler is usually more capable of reducing clutter. .

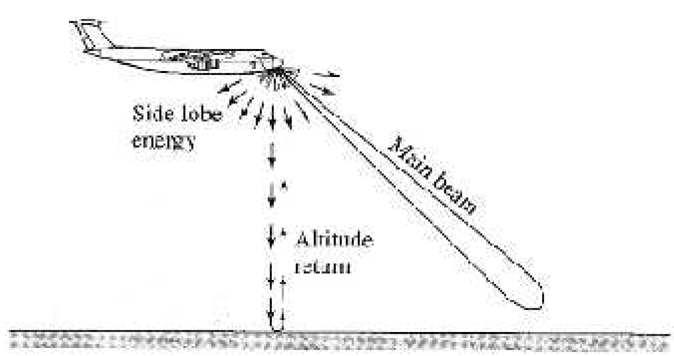


Fig. 4.9 Sketch of airborne Pulse Doppler radar

- A radar that increases its prf high enough to avoid the problems of blind speeds is called as Pulse radar.
- A high-prf pulase Doppler radar is one with no blind speeds with in the Doppler space.
- A medium-prf pulase Doppler radar is one get operated at slightly lower prf and accepts both range and Doppler ambiguities.
- A brief comparison between different Doppler pulse radar is given in table 4.1

Table. 4.1:- Comparison of different pulse Doppler radar

Radar	prf*	Duty Cycle*
X-band high-prf pulse doppler	100–300 kHz	< 0.5
X-band medium-prf pulse doppler	10–30 kHz	0.05
X-band low-prf pulse radar	1–3 kHz	0.005
UHF low-prf AMTI	300 Hz	Low

References

1- www.wikipedia.com

2- Introduction to Radar Systems by Merrill I. Skolnik, 3rd Edition, PHI Publications.

DIGITAL TELEVISION TRANSMISSION STANDARDS

A great deal of fear, uncertainty, and doubt can arise among engineers with an analog or radio-frequency (RF) background at the mere mention of digital transmission systems. Engineers sometimes fall into the trap of believing that digital systems are fundamentally different from their analog counterparts. As will be demonstrated, this is not the case. In concept, the transmission of digital television signals is no different than for analog television. The difference is in the details of implementation (hence the need for this book).

A block diagram of a typical broadcast transmission system is shown in Figure 1-1. This block diagram may, in fact, represent either an analog or a digital system. Major components include a transmitter comprising an exciter, power amplifier, and RF system components, an antenna with associated transmission line, and many receiving locations. Between the transmitter and receivers is the over-the-air broadcast transmission path. The input to the system is the baseband signal by which the RF carrier is modulated. In an analog system the baseband signal includes composite video and audio signals. In separate amplification, these modulate separate visual and aural carriers. If common amplification is used, the modulated signals are combined in the exciter and amplified together in the power amplifier. The combined signals are then transmitted together through the remainder of the link.

For a digital system, the conceptual block diagram most resembles common amplification. A single baseband signal modulates a carrier and is amplified in the transmitter, broadcast by means of the antenna, and received after propagating through the over-the-air link. The baseband signal is a composite digital data stream that may include video and audio as well as data. Since the method of modulation is also digital, the exciter used with the transmitter is also different. Beyond these details, the remainder of the system is fundamentally

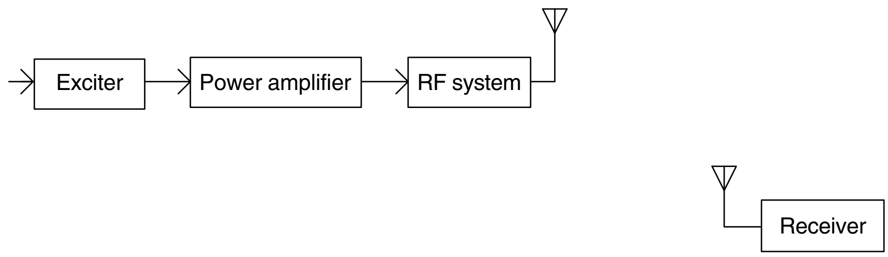


Figure 1-1. Broadcast transmission system.

the same, although there are further subtle differences in power measurement, tuning, control, and performance measurement, upconverters, power amplifiers, transmission lines, and antennas.

The similarities between digital and analog systems is also apparent when we consider the transmission channel. The ideal channel would transfer the modulated RF carrier from the modulator to the receiver with no degradation or impairment other than a reduction in the signal level and the signal-to-noise ratio. As a matter of fact, the real transmission channel is far from ideal. The signal may suffer linear and nonlinear distortions as well as other impairments in the transmitter and other parts of the channel. For analog television signals, these impairments are characterized in terms of noise, frequency response, group delay, luminance nonlinearity, differential gain, incidental carrier phase modulation (ICPM), differential phase, lower sideband reinsertion, and intermodulation distortion. For digital signals, linear distortions are also characterized in terms of frequency response and group delay. For nonlinear distortions, AM-to-AM and AM-to-PM conversion are the operative terms. In either case, the objective of good system design is to reduce these distortions to specified levels so that the channel may be as transparent as possible.

The antenna and transmission line may introduce some of the linear distortions. In most cases, these are relatively small compared to distortions introduced by the propagation path. This is especially true of matched coaxial transmission lines. Waveguides may introduce nontrivial amounts of group delay. Under some circumstances an antenna may introduce significant frequency response, nonlinear phase, and group delay distortion. Once the system design is finalized, however, no attempt may be made to equalize distortions introduced by the transmission line or antenna.

The propagation path from the broadcast antenna to the receiver location may be the source of the most significant impairments. These impairments include noise and linear distortions resulting from reflections and other sources of multipath. Depending on specific site characteristics, the linear distortions may be severe. The impairments introduced by propagation effects vary from location to location and are also a function of time. Obviously, there is no practical means of equalizing these distortions at the transmitter. Any equalization to mitigate response and group delay introduced by the over-the-air path must be done in the receiver. The random noise introduced in the propagation path may be overcome at the transmitter only by increasing the average effective radiated power (AERP).

ATSC TERRESTRIAL TRANSMISSION STANDARD

At the time of this writing, the U.S. Federal Communications Commission (FCC), Canada, and South Korea have adopted the standard developed for digital television by the Advanced Television Systems Committee (ATSC). This standard, designated A/53, represents the results of several years of design, analysis, testing, and evaluation by many experts in industry and government. It promises to be a sound vehicle for digital television delivery for decades to come. The standard describes the system characteristics of the U.S. digital television system, referred to in this book as the ATSC or DTV system. The standard addresses a wide variety of subsystems required for originating, encoding, transporting, transmitting, and receiving of video, audio, and data by over-the-air broadcast and cable systems. The transmission system is a primary subject of this book, which is described in detail in Appendix D of the ATSC standard. The ATSC standard specifies a system designed to transmit high-quality digital video, digital audio, and data over existing 6-MHz channels. The system is designed to deliver digital information at a rate of 19.29 megabits per second (Mb/s).

The transmitter component affected most by the implementation of this standard is the exciter, although, only portions of the exciter need be affected. Figure 1-2 is a conceptual block diagram of a television exciter. As drawn, this block diagram could represent either an analog or a digital exciter. The first block, the modulator, represents composite video and audio processing and modulation in the case of analog television; for digital television, this block represents digital data processing or channel coding and modulation. (It is assumed that the reader is familiar with analog video and audio modulator functions; if not, refer to Chapter 6.2, "Television Transmitters," of the *NAB Engineering Handbook*, 9th edition.)

The second block, intermediate frequency (IF)-to-RF conversion, represents upconversion, IF precorrection and equalization, final amplification, and filtering. In principle, this block is the same for both analog and digital television signals in that the main purpose is to translate the IF to the desired RF channel. For the time being, the discussion will focus on processing the digital baseband signal prior to upconversion. To facilitate this, the nature of the input and output signals of the digital modulator block is first discussed.

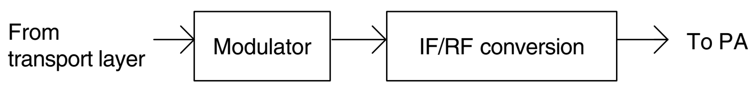


Figure 1-2. Block diagram of TV exciter.

The digital input signal to the ATSC transmission system is a synchronous serial MPEG[1]-2 transport stream at a constant data rate of 19.39... Mb/s. This serial data stream is comprised of 187-byte MPEG data packets plus a sync byte. The payload data rate is 19.2895. Mb/s. The payload may include encoded packets of digital video, digital audio, and/or data. The transport stream arrives at the exciter input on a single 75-Ω coaxial cable with a BNC input connector. The data clock is embedded with the payload data. Biphase mark coding is used. The data clock frequency error is specified to be less than ±54 Hz. The standard input level is 0.8 V ± 10% peak to peak as defined by the SMPTE Standard 310M-Synchronous Serial Interface for an MPEG-2 digital transport stream.

The output signal from the modulator block is an eight-level vestigial sideband modulated signal. Ordinarily, this is at some frequency intermediate to the baseband and RF channel frequency. The frequency, level, and other interface characteristics of the IF are generally dependent on the design choices made by the equipment manufacturer.

Figure 1-3 is a simplified block diagram of the signal processing functions required to convert the MPEG-2 transport stream to the eight-level vestigial sideband signal (8 VSB) required by the ATSC transmission system. The modulator may be viewed as performing two essential functions. The first function is channel coding. Among other things, the channel coder modifies the input data stream from the transport layer by adding information by which the receiver may detect and correct transmission errors. These are errors as a result of impairments introduced in the transmission channel. Without channel coding, the receiver would be unable to decode and display the signal properly except at receive sites with a very high signal-to-noise ratio and a minimum of multipath. The second block in Figure 1-3 is the modulator proper. It is in this block that an IF signal is modulated with the channel-coded data stream to produce the 8 VSB signal required for terrestrial over-the-air transmission.

A block diagram of the channel coder is shown in Figure 1-4. Six major functions are performed in the channel coder: data randomizing, Reed–Solomon (R/S) coding, data interleaving, trellis coding, sync insertion, and pilot signal insertion.

The incoming data from the transport stream are first randomized. This process exclusive-ORs the data bytes with a pseudorandom binary sequence locked to the data frame. The purpose of randomization is to assure that the data spectrum is uniform throughout the 6-MHz channel, even when the data are constant.

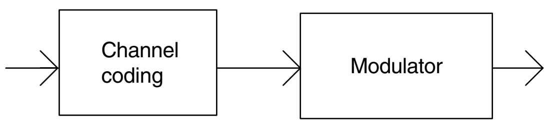


Figure 1-3. DTV modulator.

[1] Motion Pictures Expert Group.

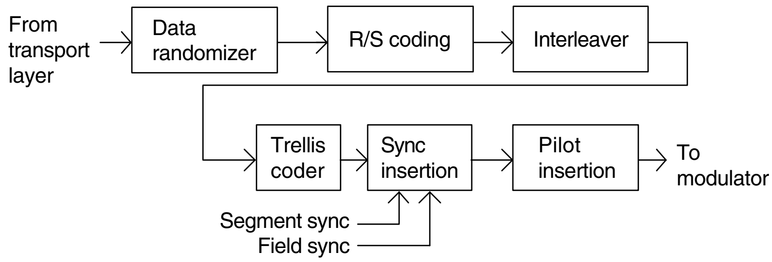


Figure 1-4. DTV channel coding. (From ATSC DTV Standard A/53, Annex D; used with permission.)

This pseudorandom sequence is generated in a 16-bit shift register with nine feedback taps. A complementary derandomizer is provided in the receiver to recover the original data sequence. Randomizing is not applied to the sync byte of the transport packet.

The next step is R/S coding. This is a forward error correction (FEC) code designed to protect against noise bursts. In this code, 20 parity bytes are added to each data block or 187-byte data packet. The R/S code selected is capable of correcting up to 10-byte errors per data block. Because of the additional bytes, the clock and data rate is necessarily increased from 19.39 Mb/s to 21.52 Mb/s. As with randomization, R/S coding is not applied to the sync bytes.

After R/S coding, the data structure is formatted into data bytes and segments, fields, and frames as defined in Figure 1-5. A data field is comprised of 312 data segments plus a sync segment, for a total of 313 segments. A data frame is comprised of two data fields, or 626 segments. The R/S coded data are interleaved to provide additional error correction. This process spreads the data bytes from several R/S packets over a much longer period of time so that a very long burst of noise is required to overrun the capability of the R/S code. A total of 87 R/S packets are processed in the interleaver.

Trellis coding, another error correction code, follows the R/S interleaver. The purpose of this code differs from the R/S code in that it has the effect of improving the signal-to-noise ratio (S/N) threshold in the presence of thermal or white noise. It is termed a $\frac{2}{3}$-rate code because every other input bit is encoded to 2 output bits; the alternate bit is not encoded. Thus the output of the trellis coder is a parallel bus of 3 bits for every 2 input bits. The trellis-coded data are interleaved with a 12-symbol code interleaver. The data rate at the output of the trellis coder is increased by a ratio of $\frac{3}{2}$, to 32.28 Mb/s. Taken together, the output bits of the trellis coder comprise the 3-bit symbols. These symbols (-7, -5, -3, -1, 1, 3, 5, 7) are the eight levels of the VSB modulator. The symbol rate is one-third that of the trellis-coded data rate, or 10.76 symbols/s.

The spectral efficiency, η_s, is the ratio of the encoded data rate to the channel bandwidth:

$$\eta_s = \frac{32.28}{6} = 5.38 \text{ bps/Hz}$$

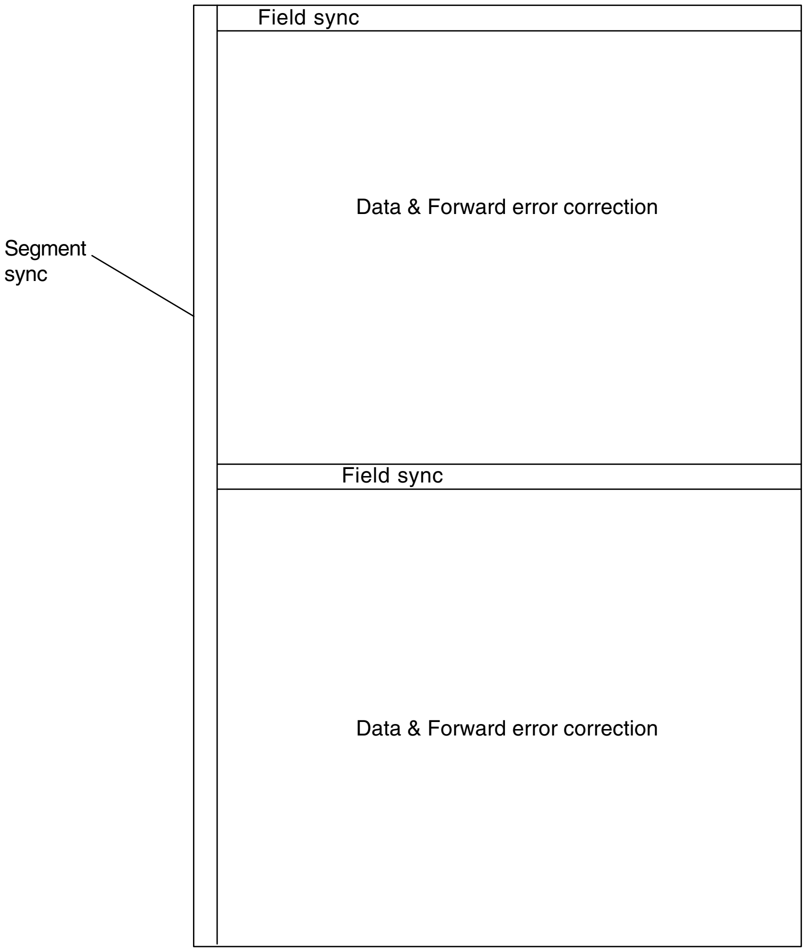


Figure 1-5. Data frame structure for the ATSC system. (From ATSC DTV Standard A/53, Annex D; used with permission.)

This is a consequence of using 3 bits per symbol to create the eight VSB levels ($M = 8$) and the excess bandwidth of the Nyquist filter ($\alpha_N = 0.1152$). Using these parameters, the spectral efficiency may be computed by

$$\eta_s = \frac{2 \ \log_2 M}{1 + \alpha_N} \qquad \text{bps/Hz}$$

which also results in 5.38 bps/Hz.

A data segment is comprised of the equivalent of the data from one R/S transport packet plus FEC code and data segment sync as shown in Figure 1-6. Actually, the data come from several R/S packets because of interleaving.

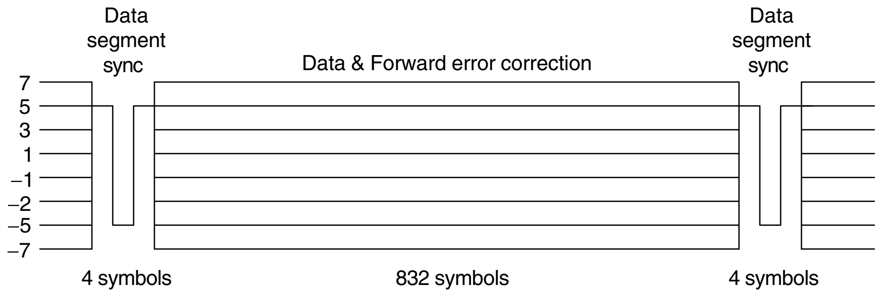


Figure 1-6. Data segment for the ATSC system. (From ATSC DTV Standard A/53, Annex D; used with permission.)

Since each R/S packet is 207 bytes in length, a data segment is 208 bytes $(207 + 1)$. At 8 bits per byte and 3 bits per symbol, the data segment is 208 bytes $\times$ 8 bits/byte $\times$ $\frac{3}{2}/3$ bits/symbol, or 832 symbols in length, 828 of which are FEC coded data; the remaining four are segment sync symbols. There are 3×832, or 2496 bits per segment, 2484 of which are data and 12 of which are segment syncs. For the data rate of 32.28 Mb/s the time per bit is 31 ns. Thus the time per segment is 2496×31, or 77.3 μs, and the segment rate, $f_{\text{seg}} = 12.94$ data segments per second. With 313 segments per field, the field time is 313×77.3 μs, or 24.2 ms, and the field rate is 41.3 kHz. The frame rate, f_{frame}, is one-half the field rate, or 20.66 kHz.

Following the trellis coding, field and segment sync symbols are inserted. The structure of the data field sync segment is defined in Figure 1-7. As with the data segments, the field sync segment is 832 symbols in length. Each symbol is binary encoded as either $+$ or -5. Four data segment sync symbols replace the MPEG sync byte. These are followed by a series of pseudorandom number (PN) sequences of length 511, 63, 63, and 63 symbols, respectively. The PN63 sequences are identical, except that the middle sequence is of opposite sign in every other field. This inversion allows the receiver to recognize the alternate data fields comprising a frame.

The PN63 sequences are followed by a level identification sequence consisting of 24 symbols. The last 104 symbols of the field sync segment are reserved; 92 of these symbols may be a continuation of the PN63 sequence. The last 12 of these symbols are duplicates of the last 12 symbols of the preceding data segment.

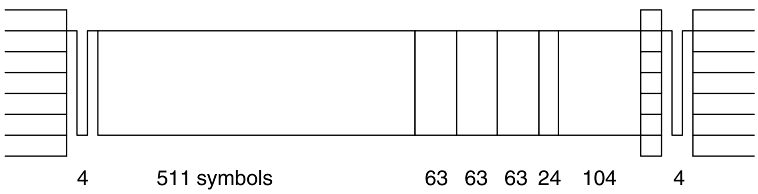


Figure 1-7. Field sync for the ATSC system. (From ATSC DTV Standard A/53, Annex D; used with permission.)

In addition to providing a means of synchronizing the receiver to the formatted data, the sync segments serve as training signals for the receiver equalizer. The equalizer improves the quality of the received signal by reducing linear distortions. This is analogous to ghost reduction due to multipath in analog systems. Since the sync sequences are known repetitive signals, the equalizer taps may be adjusted to reproduce these sequences with a minimum of distortion. The taps, thus adjusted, reduce distortion of the received data. The sync segments may also be used for diagnostic purposes.

The data field and frame structure has the familiar appearance of the field and frame structure of analog television. However, it should not be assumed that a data field corresponds to a video field. Each data field may include video, audio, or other data, so there is generally no correspondence between data fields and video fields.

VESTIGIAL SIDEBAND MODULATION

Vestigial sideband modulation may be accomplished in either the analog or the digital domain. Manufacturers have generally developed their own modulation schemes, some of which may be proprietary. Since the purpose of this book is to describe the principles of digital television transmission, a generic modulator using analog circuitry is presented.

Such a modulator is illustrated in Figure 1-8. The signal (i.e., the 3-bit multilevel symbols or pulses from the output of the trellis coder) is divided equally to form in-phase (I) and quadrature (Q) paths at the input to the modulator. The pulses are then shaped to minimize intersymbol interference. This pulse shaping is accomplished in a Nyquist filter. This is a low-pass linear-phase filter with flat amplitude response over most of its passband. At the upper and lower band edges, the filter response transitions to the stopband by means of skirts with a root-raised-cosine shape. The steepness of the skirts is determined by the shape factor, α_N. For the ATSC system, α_N is specified to be 0.1152. The Nyquist filter multiplies the shaped signals by either $\sin(\pi t/2T)$ or $\cos(\pi t/2T)$, where T is the symbol time.

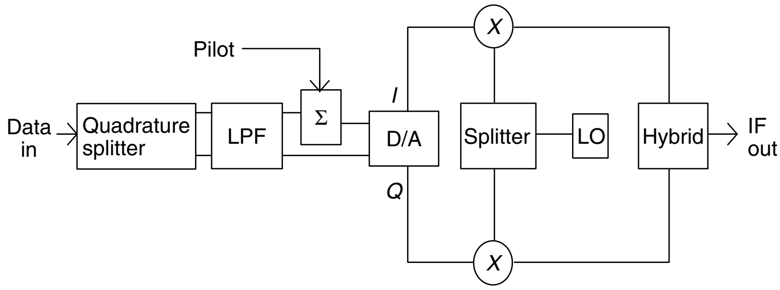


Figure 1-8. Typical digital modulator.

The shaped I and Q signals are now presented to digital-to-analog (D/A) converters in each of the I and Q channels. The I and Q signals are each multiplied by equal levels of the local oscillator (LO) signal. For the Q path, the LO signal is 90° out of phase with respect to the LO signal for the I path. These signals are then summed in a two-way power combiner to produce the IF output. The resulting spectrum contains only one of the sidebands of the modulated signals and the carrier is suppressed. Thus this modulation technique is called vestigial sideband. A pilot signal is inserted in the I path of the modulator. By adding a small direct-current (dc) offset of 1.25 V to all of the encoded symbols (including sync), a tone at the same frequency as the suppressed carrier is generated in the output of the VSB modulator. The presence of the pilot adds very little power (only 0.3 dB) to the modulated signal, but it is important in that it enables receiver tuning under conditions of severe noise and interference. It also speeds carrier recovery and, therefore, data acquisition in the receiver. It is apparent that the quality of the IF output is dependent on the stability of both the incoming data and the LO.

At this point in the system, the complete DTV signal has been generated, consisting of eight amplitude levels, four positive and four negative. The signal is often displayed in a two-dimensional I–Q or constellation diagram, as shown in Figure 1-9. This is a graphical representation of the orthogonal I and Q components of the modulated waveform, plotted in X–Y or rectangular

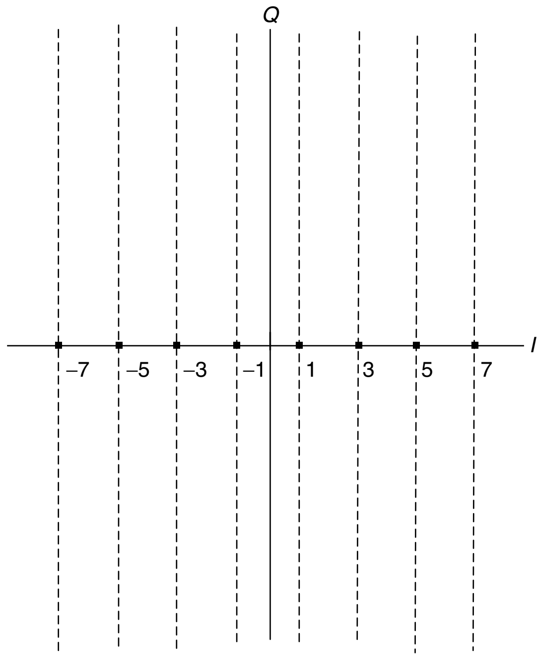


Figure 1-9. I–Q diagram for 8 VSB signal.

coordinates, where the X and Y axes are called the I and Q axes, respectively. Each point in the I–Q diagram represents a specific amplitude and phase of the RF carrier. For 8 VSB, information is carried only by the I component, for which the distinct levels of 8 VSB are plotted on the horizontal axis. Although a quadrature component is present and is displayed in the direction of the Q-axis, there are no distinct levels associated with the Q component and no information conveyed.

The modulated signal occupies 6 MHz of total bandwidth by virtue of the vestigial sideband modulation scheme. The spectrum of the modulated signal is shown in Figure 1-10. The energy is spread uniformly throughout most of the channel. At both the upper and lower band edges, the spectrum is shaped in accordance with the root-raised-cosine or Nyquist filter. A complementary root-raised-cosine filter of the same shape is included in the receiver so that the system response is a raised cosine function.

The 3-dB bandwidth of the resulting transmitted spectrum is 5.38 MHz. At the RF channel frequency, the pilot is located at the lower 3-dB point, 0.31 MHz above the lower edge of the channel. The pilot is the same frequency as the suppressed carrier. (The pilot may be at the opposite end of the spectrum at the IF.) The DTV pilot is offset from the NTSC visual carrier to minimize DTV-to-NTSC cochannel interference. The remainder of the system exists for the purposes of upconverting to the desired channel, amplifying to the required power level, and radiating the on-channel signal.

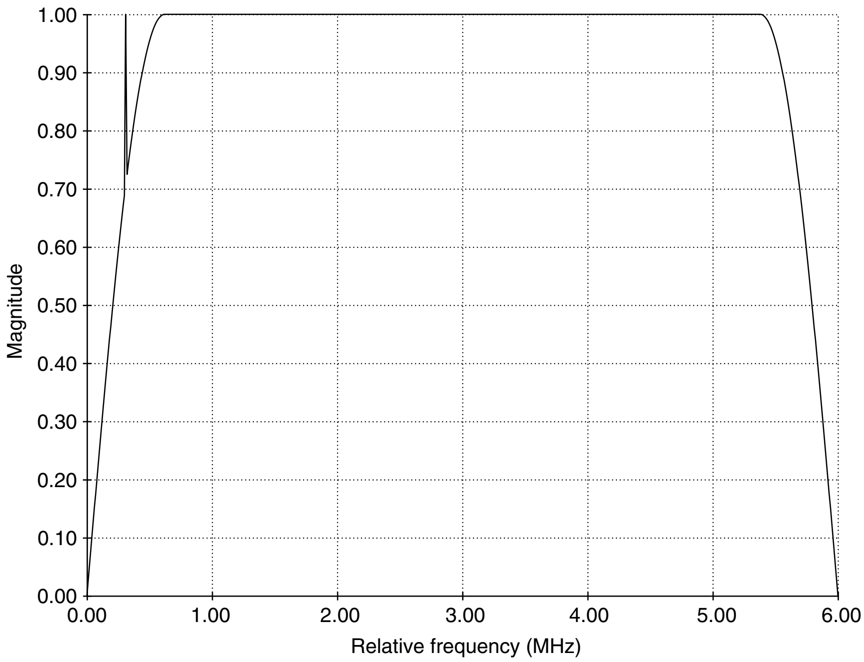


Figure 1-10. Transmitted spectrum, 8 VSB.

DVB-T TRANSMISSION STANDARD

The European Telecommunications Standards Institute has adopted a set of standards for digital broadcasting of television, sound, and data services. Standards have been adopted for satellite, cable, and terrestrial signal delivery. The standard for terrestrial transmission, ETS 300 744, is designated Digital Video Broadcast−Terrestrial (DVB-T). This standard describes a baseline transmission system for digital broadcasting of television. At the time of this writing, it has been adopted by the 15 members of the European Union, Australia, and New Zealand. It is similar in many respects to the U.S. DTV standard. However, there are also important and significant differences in both channel coding and modulation.

The DVB-T standard specifies a system designed to transmit high-quality digital video, digital audio, and data over existing 7- or 8-MHz channels. The system is designed to deliver digital information at rates from 4.98 to 31.67 Mb/s. Although there are many similarities with the ATSC standard in the transport layer and channel coding, a significant difference is in the type of modulation used. Coded orthogonal frequency-division multiplex (COFDM) has been selected for DVB-T, in part due to the unique requirements of European broadcasting stations and networks. Single-frequency networks (SFN) are used extensively in Europe to more effectively use the channels available; COFDM is seen as best suited to this requirement. In a SFN, all stations broadcasting a particular program do so on the same channel, each being synchronized to precisely the same reference signal and having common baseband timing. A receiver tuned to this channel may receive signals from one or more stations simultaneously, each with a different delay. Under multipath conditions, the signal strength from each station may vary with time. The guard intervals and equalization built into the COFDM system facilitate effective reception under these conditions. The guard interval may be selected from $\frac{1}{32}$ to $\frac{1}{4}$ the duration of the active symbol time, so that the total symbol duration is from $1\frac{1}{32}$ to $1\frac{1}{4}$ the active symbol time.

As with the ATSC standard, the transmitter assembly most affected by the transition to digital broadcast is the exciter, with the major changes required being baseband processing and modulation. Thus the focus of this discussion is on the modulator block. The nature of the input and output signals is discussed first.

In common with DTV in the United States, the digital input signal to the DVB-T transmission system is a MPEG-2 synchronous transport stream comprised of 187-byte MPEG data packets plus a sync byte. The payload may include encoded packets of digital video, digital audio, and/or data. The parallel transport stream connector at the modulator input is a DB25 female connector. The data clock line is separate from the payload data lines.

The output signal from the modulator block is a COFDM signal. Ordinarily, this is generated at some frequency intermediate to the baseband and RF channel frequency. The frequency, signal level, and other interface characteristics of the IF are generally dependent on design choices made by the equipment manufacturer.

As in the ATSC system, the modulator may be viewed as performing the functions of channel coding and modulation proper. The functions performed in the channel coder include energy dispersal or data randomization, outer or R/S coding, outer interleaving, inner or trellis coding, and interleaving. The modulator functions include mapping, frame adaptation, and pilot insertion.

The incoming data from the transport stream are first dispersed or randomized. A complementary derandomizer is provided in the receiver to recover the original data sequence. As with the DTV system, randomizing is not applied to the sync byte of the transport packet.

The next step is outer or R/S coding. The details of the code selected differ from those of the DTV standard in that it is capable of correcting up to only eight byte errors per data block. In this code, 16 parity bytes are added to each sync and data block or 188-byte packet. The R/S coded data are interleaved to provide additional error correction.

Convolutional coding and interleaving follow the R/S interleaver. The DVB-T system allows for a range of punctured convolutional codes. Selection of the code rate is based on the most appropriate level of error correction for a given service and data rate. Punctured rates of $\frac{2}{3}$, $\frac{3}{4}$, $\frac{5}{6}$, or $\frac{7}{8}$ are derived from the $\frac{1}{2}$-rate mother code. Interleaving consists of both bitwise and symbol interleaving. Bit interleaving is performed only on the useful data.

The purpose of the symbol interleaver is to map bits on to the active OFDM carriers. Detailed operation of the symbol interleaver depends on the number of carriers generated, whether 2048 (2^{11}) in the 2k mode or 8192 (2^{13}) in the 8k mode. Some of the carriers are used to transmit reference information for signaling purposes (i.e., to select the parameters related to the transmission mode). The number of carriers available for data transmission is 1705 in the 2k mode or 6817 in the 8k mode. The overall bit rate available for data transmission is not dependent on the mode but on the choice of modulation used to map data on to each carrier.

The OFDM modulator follows the inner coding and interleaving. This involves computing an inverse discrete fourier transform (IDFT) to generate multiple carriers and quadrature modulation. The transmitted signal is organized in frames, each frame having a duration of T_F and consisting 68 OFDM symbols. The symbols are numbered from 0 to 67, each containing data and reference information. In addition, an OFDM frame contains pilot cells and transmission parameter signaling (TPS) carriers. The pilot signals may be used for frame, frequency, and time synchronization, channel estimation, and transmission mode identification. TPS is used to select the parameters related to channel coding and modulation.

The many separately modulated carriers may employ any one of three square constellation patterns: quadrature-phase shift keying (QPSK, 2 bits per symbol), 16-constellation-point quadrature amplitude modulation (16 QAM, 4 bits per symbol), or 64 QAM (6 bits per symbol). By selecting different levels of QAM in conjunction with different inner code rates and guard interval ratios, bit rate may be traded for ruggedness. For example, QPSK with a code rate of $\frac{1}{2}$ and a guard interval ratio of $\frac{1}{4}$ is much more rugged than 64 QAM with a code rate of $\frac{5}{6}$ and $\frac{1}{32}$ guard interval ratio. However, the available data rate is much less. The 8k mode has the longest available guard interval, making it the best choice for single-frequency networks with widely separated transmitters.

Hierarchical transmission is also a feature of the DVB-T standard. The incoming data stream is divided into two separate streams, a low- and a high-priority stream, each of which may be transmitted with different channel coding and with different modulation on the subcarriers. This allows the broadcaster to make different trade-offs of bit rate and ruggedness for the two streams.

The power spectral density of the modulated carriers is the sum of the power spectral density of the individual carriers. The upper portion of the transmitted spectrum for an 8-MHz channel is shown in Figure 1-11, plotted relative to the channel center frequency. Overall, the energy is spread nearly uniformly throughout most of the channel. However, since the symbol time is larger than the inverse of the carrier spacing, the spectral density is not constant. The center frequency of the DVB-T channels is the same as the current European analog ultrahigh frequency (UHF) channels. The minimum carrier-to-noise (C/N) ratio is dependent, among other parameters, on modulation and inner code rate. As with the DTV system, there is no visual, chroma, or aural carrier frequencies as in analog TV.

The complete DVB-T signal has been generated at the output of the modulator. The remainder of a transmitting system exists for the purposes of upconverting

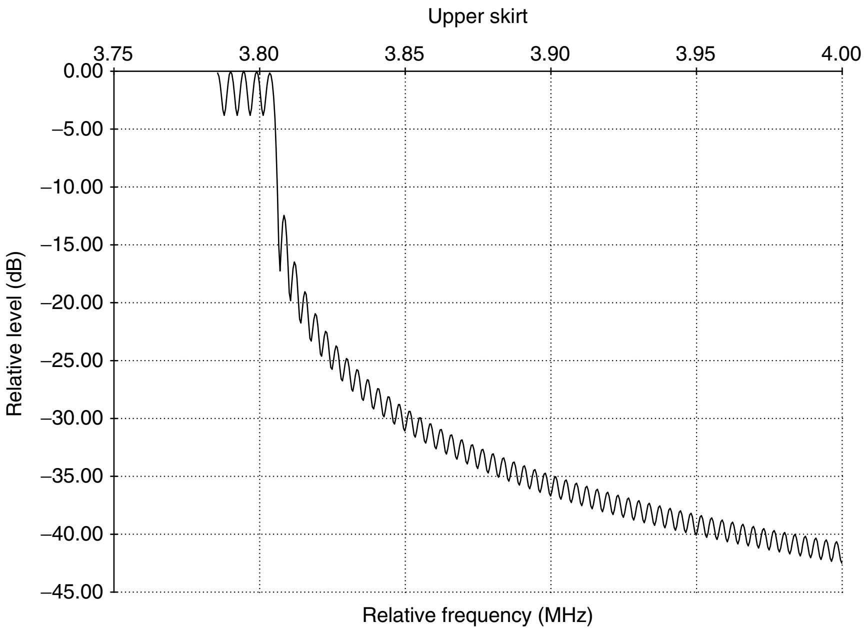


Figure 1-11. Typical COFDM spectrum.

to the desired channel, amplifying to the required power level, and radiating the on-channel signal.

ISDB-T TRANSMISSION STANDARD

Japan's Digital Broadcasting Experts Group (DiBEG) has developed a standard for digital broadcasting of television, sound, and data services, designated integrated services digital broadcasting (ISDB). Standards have been developed for delivery of satellite, cable, and terrestrial signals. These standards include a description of a baseline transmission system that provides for digital broadcasting of television, including channel coding and modulation. The transmission standard for terrestrial digital television is similar in many respects to the DVB-T standard. It is entitled Integrated Services Digital Broadcasting–Terrestrial (ISDB-T).[2] A key difference with respect to DVB-T is the use of band-segmented transmission–OFDM (BST-OFDM). This is a data segmentation approach that permits the service bandwidth to be allocated to various services, including data, radio, standard definition television (SDTV), and high-definition television (HDTV) in a flexible manner. It is planned that digital television will be launched in Japan after 2003.

The ISDB-T standard specifies a system designed to transmit over existing 6-, 7-, or 8-MHz channels. The system is designed to deliver digital information at data rates from 3.561 to 30.980 Mb/s.

In common with the other world standards, the digital input signal to the ISDB-T transmission system is a MPEG-2 synchronous transport comprised of 187-byte MPEG data packets plus a sync byte. The payload may include encoded packets of digital video, digital audio, text, graphics, and data. In addition, transmission and multiplex control (TMCC) is defined for hierarchical transmission. To make use of the band-segmenting feature, the data stream is remultiplexed and arranged into data groups, each representing all or part of a program or service. After channel coding, these data groups become OFDM segments. Each OFDM segment occupies $\frac{1}{14}$ of the channel bandwidth. This arrangement allows for both broadband and narrowband services.

For example, a single HDTV service might occupy 12 of the OFDM segments, with the thirteenth used for sound and data.[3] Alternatively, multiple SDTV programs might occupy the 12 OFDM segments. A maximum of three OFDM segment groups or hierarchical layers may be accommodated at one time. For the narrowband services, a small, less expensive narrowband receiver may be used. The OFDM segment in the center of the channel is dedicated to such narrowband or partial reception services. Obviously, a receiver decoding a single OFDM segment receives only a portion of the original transport stream.

[2] "Channel Coding, Frame Structure, and Modulation Scheme for Terrestrial Integrated Services Digital Broadcasting (ISDB-T)," *ITU Document 11A/Jxx-E*, March 30, 1999.

[3] The upper and lower channel edges occupy the bandwidth of the remaining OFDM segment.

ISDB-T has many features in common with DVB-T. Both inner and outer FEC codes are applied to the data. The resulting data stream modulates multiple orthogonal carriers. Thus both standards make use of COFDM. The guard interval may be selected from $\frac{1}{32}$ to $\frac{1}{4}$ the duration of the active symbol time. As in Europe, Japan will use this approach to increase the number of available channels by means of SFNs. The R/S code is capable of correcting up to eight-byte errors per data block. A total of 16 parity bytes are added to each sync and data block. The system also allows for a range of punctured convolutional codes. Selection of the code rate is based on the most appropriate level of error correction for a given service or data rate. Code rates of $\frac{2}{3}$, $\frac{3}{4}$, $\frac{5}{6}$, or $\frac{7}{8}$ are derived from a $\frac{1}{2}$-rate mother code.

Despite similarities, there are differences in implementation of the channel coding. As shown in Figure 1-12, the order of R/S coding and energy dispersal are interchanged with the order used in either the DTV and DVB-T systems. The R/S coding is applied to the data as they emerge from the remultiplexed transport stream, including any null packets. Energy dispersal, delay adjustment, bytewise interleaving, and trellis coding are then applied in that order to each data group separately. This permits the length of the interleaving, code rate for the inner FEC code, and signal constellation to be selected independently for each hierarchical layer. The null packets at the output of the R/S coder are removed. The delay resulting from the bytewise interleaving differs for each layer, depending on the channel coding and modulation. To compensate, a delay adjustment is inserted prior to the interleaver.

The OFDM modulator follows the inner coding and interleaving. This involves computing an IDFT to generate multiple carriers and quadrature modulation. The number of carriers available ranges from 1405 to 5617 for all channel bandwidths depending on the transmission mode. Of these, the number of carriers available for data transmission ranges from 1249 to 4993. Obviously, the carrier spacing is increased for the wider channels for a given number of carriers. The information bandwidth is approximately 5.6, 6.5, and 7.4 MHz for the 6-, 7-, and 8-MHz channels, respectively.

The transmitted signal is organized in frames. However, the frame duration is not fixed as in the DVB-T standard. Rather, the frame duration depends on the

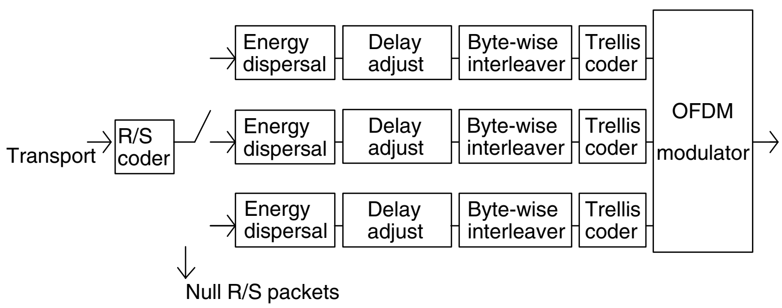


Figure 1-12. Block diagram of ISDB-T channel coding.

TABLE 1-1. Frame Duration (ms) versus Mode and Guard Interval Ratio

Guard Interval Ratio	Mode		
	1	2	3
$\frac{1}{4}$	64.26	128.52	257.04
$\frac{1}{8}$	57.83	115.67	231.34
$\frac{1}{16}$	54.62	109.24	218.46
$\frac{1}{32}$	53.01	106.03	212.06

transmission mode and the length of the guard interval. These relationships are summarized in Table 1-1. The frame duration doubles from mode 1 to mode 2 and doubles again from mode 2 to mode 3. Modes 1, 2, and 3 are defined for 108, 216, and 432 OFDM carriers per segment, respectively. The packets comprising the frames are numbered consecutively and contain the payload data as well as the information necessary for operation of the broadcast system. Scattered and continual pilot signals are available for frequency synchronization and channel estimation. The TMCC subchannels carry packets with information on the transmission parameters. Auxiliary subchannels carry ancillary information for network operation.

The many separately modulated carriers may employ the same three constellation patterns as provided in the DVB-T standard-QPSK, 16 QAM or 64 QAM. In addition, differential quadrature-phase shift keying (DQPSK) is available.

With modulation complete, the complete ISDB-T signal has been generated. The remainder of the transmitting system exists for the purposes of upconverting to the desired channel, amplifying to the required power level, and radiating the on-channel signal.

CHANNEL ALLOCATIONS

The DVB-T system operates in 7- and 8-MHz channels primarily within the existing European UHF spectrum (bands IV and V), although implementation guidelines have been published for bands I and III as well.[4] In Japan, 6-MHz channels are to be used. In general, the availability of spectrum varies from country to country; in virtually every case a scarcity exists. As in the United States, the tendency is to move terrestrial television to the UHF spectrum to free other frequencies for other uses. A common solution in Europe, Japan, and other countries is to use regional and/or national SFNs. This approach allows for the broadcast of just a few programs to a high percentage of the target area using a minimum number of channels. Generally, the digital services will coexist

[4] *Implementation Guidelines for DVB-T: Transmission Aspects*, European Telecommunications Standards Institute, April 1997.

with existing analog services for some extended period of time (say, 15 years), after which the analog service will be discontinued. Where possible, existing transmitter sites, antennas, and towers are expected to be used. To maximize the possibility of viewers using existing receiving antennas, assignment of channels near the existing analog channels with the same polarization is desirable.

In the United States, the DTV system operates in 6-MHz channels in portions of both the very high frequency (VHF) and UHF spectrum. A core spectrum is defined which includes a total bandwidth of 294 MHz, extending from channel 2 through channel 51. Because of the limited availability of spectrum and the need to minimize interference, some broadcasters are assigned spectrum outside the core at frequencies as high as channel 69 during the transition period. Use of channel 6 is minimized due to the potential for interference with the lower FM frequencies. The use of Channels 3 and 4 in the same market is minimized to facilitate use of cable terminal devices, which may operate on either of these channels. Channel 37 is reserved for radio astronomy. Where it is necessary to use adjacent channels in the same market, the NTSC and DTV stations are colocated, if possible. The licensees for the adjacent channels are required to lock the DTV and NTSC carrier frequencies to a common reference frequency to protect the NTSC from excessive interference.

The U.S. Telecommunications Act of 1996 provides that initial eligibility for an advanced television license is limited to existing broadcasters with the condition that they eventually relinquish either the current analog channel or the new digital channel at the end of the transition period. The purpose of this provision is, in part, to promote spectrum efficiency and rapid recovery of spectrum for other purposes. Consequently, DTV was introduced in the United States by assigning existing broadcasters with a temporary channel on which to operate a DTV station during the transition period, which will extend to 2006. It is planned that 78-MHz of spectrum will be recovered at the end of the transition period; it is also planned that 60 MHz in channels 60 to 69 will be recovered earlier.

If in the future channels 2 to 6 prove to be acceptable for transmission of DTV, the core spectrum may be redefined to be channels 2 to 46.[5] Those stations operating outside the core spectrum during the transition will be required to move their DTV operations to a channel inside the core when one becomes available. Broadcasters whose existing NTSC channel is in the core spectrum could move their DTV operations to this channel in the future. Broadcasters whose NTSC and DTV channels are in the core spectrum could chose which of those will be their permanent DTV channel.

It is evident that broadcast engineers in the United States will face the challenge of transmission system design and operation for the foreseeable future. Many important decisions must be made for initial systems during the transition period. Many of these systems will continue to operate without the need for major changes after the transition period ends. Others systems, however, may require major changes to accommodate the channel shifts required by the FCC.

[5] *FCC 6th Report and Order DTV Allocations*, Appendix D, April 22, 1997, p. D-11.

ANTENNA HEIGHT AND POWER

In the United States, the antenna height above average terrain (HAAT) and AERP for DTV stations operated by existing licensees is designed to provide equivalent noise-limited coverage to a distance equal to the present NTSC grade B service contour. The maximum permissible power for new DTV stations in the UHF band is 316 kW. The maximum antenna height is 2000 ft above average terrain. For HAATs below this value, higher AERP is permitted to achieve equivalent coverage. The maximum AERP is 1000 kW regardless of HAAT. The minimum AERP for UHF is 50 kW. Power allocations for VHF range from 200 W to slightly more than 20 kW.

MPEG-2

Although the source encoding and transport layer are distinct from the transmission system, they are closely associated. It is therefore important that the transmission system engineer have an understanding of MPEG-2. The following discussion is a cursory overview; for more details, the interested reader is referred to ATSC A/53 or the Implementation Guidelines for DVB-T, which point to additional documents.

In accordance with the International Telecommunications Union, Radio Sector (ITU-R) digital terrestrial broadcast model, the transport layer supplies the data stream to the RF/transmission system. This is illustrated in Figure 1-13. Since there is no error protection in the transport stream, compatible forward error correction codes are supplied in the transmission layer as already described.

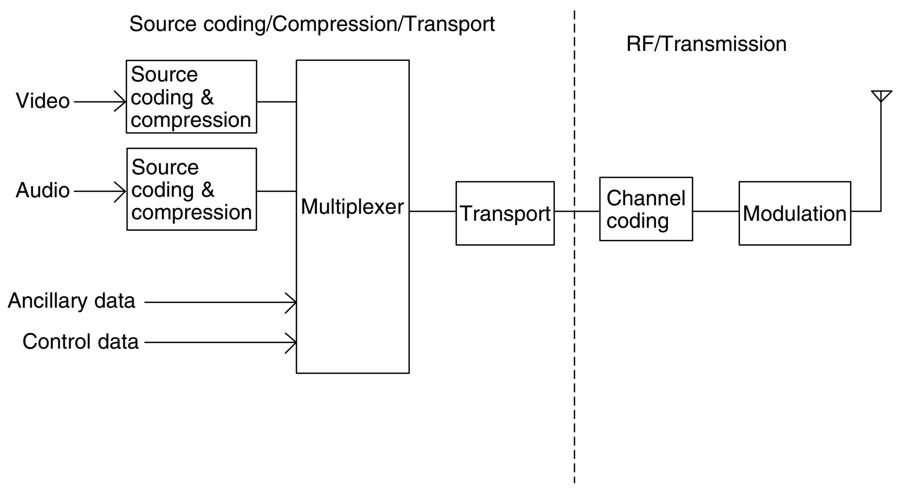


Figure 1-13. Digital television broadcast model. (From ATSC DTV Standard A/53, Annex D; used with permission.)

MPEG-2 refers to a set of four standards adopted by the International Standards Organization (ISO). Together, these standards define the syntax for the source coding of video and audio and the packetization and multiplexing of video, audio, and data signals for the DTV, DVB-T, and ISDB-T systems. MPEG-2 defines the protocols for digital compression of the video and audio data. These video coding "profiles" allow for the coding of four source formats, ranging from VCR quality to full HDTV, each profile requiring progressively higher bit rates. Several compression tools are also available, each higher level being of increased sophistication. The sophistication of each level affects the video quality and receiver complexity for a given bit rate. In general, the higher the bit rate, the higher the video and audio quality. Tests indicate that studio-quality video can be achieved with a bit rate of about 9 Mb/s. Consumer-quality video can be achieved with a bit rate ranging from 2.5 to 6 Mb/s, depending on video content.

Audio compression takes advantage of acoustic masking of low-level sounds at nearby frequencies by coding these at low data rates. Other audio components that cannot be heard are not coded. The result is audio quality approaching that of a compact disk at a relatively low data rate. The transport format and protocol are based on a fixed-length packet defined and optimized for digital television delivery. Elementary bit streams from the audio, video, and data encoders are packetized and multiplexed to form the transport bit stream. Complementary recovery of the elementary bit streams is made at the receiver.

The transport stream is designed to accommodate a single HDTV program or several standard definition programs, depending on the broadcaster's objectives. Even in the case of HDTV, multiple data sources are multiplexed, with the multiplexing taking place at two distinct levels. This is illustrated in Figure 1-14. In the first level, program bit streams are formed by multiplexing packetized elementary streams from one or more sources. These packets may be coded video, coded audio, or data. Each of these contain timing information to assure that each is decoded in proper sequence.

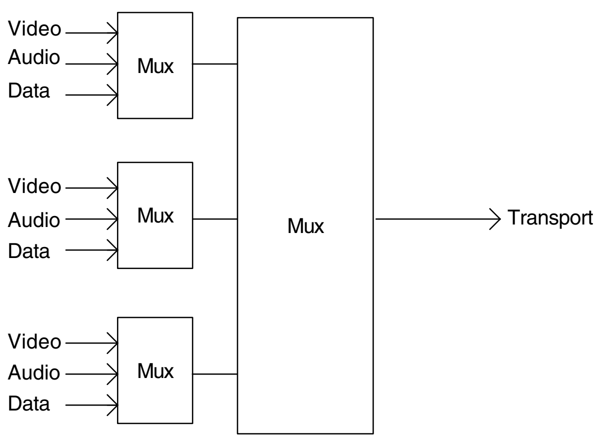


Figure 1-14. MPEG-2 multiplexing.

A typical program might include video, several audio channels, and multiple data streams. In the second level of multiplexing, many single programs are combined to form a system of programs. The content of the transport stream may be varied dynamically depending on the information content of the program sources. If the bit rate of the multiplexed packets is less than the required output bit rate, null packets are inserted so that the sum of the bit rates matches the constant bit rate output requirement. All program sources share a common clock reference. The transport stream must include information that describes the contents of the complete data stream and access control information, and may include internal communications data. Scrambling for the purpose of conditional access and teletext data may also be accommodated. An interactive program guide and certain system information may be included.

As implemented in the ATSC system, the video and audio sampling and transport encoders are frequency locked to a 27-MHz clock. The transport stream data rate and the symbol rate are related to this clock. If the studio and transmitter are colocated, the output of the transport stream may be connected directly to the transmitter. In many cases, the transport stream will be transmitted via a studio-to-transmitter link (STL) to the main transmitter site. This requires demodulation and decoding of the STL signal to recover the transport stream prior to modulation and coding in the DTV, DVB-T, or ISDB-T transmitter.

PERFORMANCE OBJECTIVES FOR DIGITAL TELEVISION

Characterization of the signal quality is an aspect in which digital systems differ most from their analog counterparts. With analog TV signals, engineers can readily measure the transmitted or received power at the peak of the sync pulse. The average power varies depending on picture content. Methods are available for separately measuring aural and chroma carrier power levels. Nonlinear distortions are characterized by differential gain and phase, luminance nonlinearity, and ICPM. Linear distortions are evaluated in terms of swept response and group delay.

For digital television systems, some of the familiar performance measurements are somewhat elusive. A regularly recurring sync pulse is not available for the purpose of measuring peak envelope power. The data representing video, chroma, and sound are multiplexed into a common digital stream; separate visual, chroma, and aural carriers do not exist. Because of the random nature of the baseband signal, the average power within the transmission bandwidth is constant. The quality measures of interest include average power, peak-to-average power ratio, carrier-to-noise ratio (C/N),[1] the ratio of the average energy per bit to the noise density (E_b/N_0), symbol and segment error rates (SER), bit error rate (BER), error vector magnitude (EVM), eye pattern opening, intersymbol interference (ISI), AM-to-AM conversion, AM-to-PM conversion, and spectral regrowth. Characterization of linear distortion by frequency response and group delay is common for both analog and digital systems.

[1] Reference to signal-to-noise ratio (S/N) and carrier-to-noise ratio (C/N) will be found in the literature with no distinction in meaning. In other works, C/N refers to predetection or input signal-to-noise power ratio, S/N to postdetection or output signal-to-noise power ratio. The latter convention is followed in this book.

Channel capacity is a function of carrier-to-noise ratio and channel bandwidth. Therefore, the factors affecting system noise and transmission errors at the receiver are discussed first. Following this is a discussion of factors that describe transmitter performance.

SYSTEM NOISE

Ideally, a digital television transmission system should provide an impairment-free signal to all receiving locations within the service area. Obviously, there will be some locations where this ideal cannot be achieved. In a practical system, linear distortions, nonlinear distortions, and various sources of noise and interference will impair the signal. The overall effect of these impairments is to degrade the carrier-to-noise plus interference ratio ($C/(N+I)$). In the absence of interference, this term reduces to the more familiar C/N.

Consider first the case for which there is no interference from other digital or analog signals. Knowing the received signal power and the noise power at the receiving location allows determination of the C/N and the noise-limited coverage contour in the absence of multipath and interference. Methods of determining the average power of the received signal, P_r, are discussed in Chapter 8. In the following discussion, the average carrier power, C, is considered to be equivalent to P_r after adjustment for receive antenna gain and downlead attenuation.

At distant receive locations, thermal noise should be the predominate noise source in the absence of severe multipath or interference. Thermal noise is often assumed to be additive white Gaussian noise (AWGN). The noise power spectrum of AWGN is flat over an infinite bandwidth with a power spectral density of $N_0/2$ watts per hertz.[2] The total noise power, N, in a channel of bandwidth, B, is the product of N_0 and B,

$$N = N_0 B$$

Much of the thermal noise power is due to the noise generated in input stages of the receiver. Total noise power at the receiver input may be expressed as

$$N = kT_sB \qquad \text{watts}$$

where k is Boltzmann's constant (1.38×10^{-23} Joules/Kelvin) and T_s is the receive system noise temperature in Kelvins. This formula may be written in terms of decibels above a milliwatt (dBm)

$$N(\text{dBm}) = -198.6 + 10 \log B + 10 \log T_s$$

[2] The assumption of white noise is not strictly true for all sources of noise. For example, noise from galactic sources decreases with increasing frequency. However, for all practical purposes over the bandwidth of one channel, the noise spectrum may be considered to be flat.

For DTV transmission in the United States, the channel bandwidth is 6 MHz, so that the thermal noise limit for a perfect receiver at room temperature, N_t, is

$$N_t = 1.38 \times 10^{-23} \times 290 \times 6 \times 10^6 = 24.01 \times 10^{-15}\ \mathrm{W}$$

Converting to dBm, the thermal noise limit is −106.2 dBm. For the 7- and 8-MHz channels provided for in the DVB-T and ISDB-T standards, the thermal noise limit is −105.7 and −105.2 dBm,[3] respectively.

To determine the threshold receiver power, P_{mr}, required at the receiver, the threshold carrier-to-noise ratio and receiver noise figure, NF, must be added to the thermal noise limit. That is,

$$P_{\mathrm{mr}} = N_t + C/N + \mathrm{NF}$$

To determine the threshold power at the antenna, the line loss ahead of the receiver must be added and the receive antenna gain subtracted from the threshold receiver power:

$$P_{\mathrm{ma}} = P_{\mathrm{mr}} - G_r + L$$

For planning purposes in the United States, the FCC Advisory Committee on Advanced Television Service has recommended standard values for receiver noise figure, the loss of the receiving antenna transmission line, and antenna gain at the geometric mean frequency of each of the RF bands.[4] These planning factors are shown in Table 2-1. The resulting threshold received power at the antenna and receiver terminals is also shown in the last two lines of this table. Satisfactory reception is defined in terms of the threshold of visibility (TOV). For the U.S. DTV system this is set at a threshold C/N value of 15.2 dB.

A similar table for the DVB-T system using 8-MHz channels is constructed in Table 2-2. For this system, the theoretical threshold C/N for nonhierarchical transmission in a Gaussian channel ranges from 3.1 to 29.6 dB.[5] For Table 2-2,

TABLE 2-1. FCC Planning Factors and Threshold Power

	VHF		
Component	Low	High	UHF
Receiver antenna gain, G_r (dB)	4	6	10
Line loss, L (dB)	1	2	4
Noise figure, NF (dB)	10	10	7
Threshold C/N (dB)	15.2	15.2	15.2
Threshold power at antenna, P_{ma} (dBm)	−84.0	−85.0	−90.0
Threshold power at receiver, P_{mr} (dBm)	−81.0	−81.0	−84.0

[3] The equivalent noise bandwidth for an 8-MHz channel is actually 7.6 Mhz.

[4] *FCC Sixth Report and Order*, April 3, 1997, p. A-1.

[5] *ETS 300 744*, March 1996, pp. 38–41.

TABLE 2-2. DVB-T Minimum Receiver Signal Input Levels for 8-MHz Channels

	Band			
Component	I	III	IV	V
Receiver antenna gain, G_r (dB)	3	7	10	12
Line loss, L (dB)	1	2	3	5
Noise figure, NF (dB)	5	5	5	5
Threshold C/N (dB)	13.9	13.9	13.9	13.9
Threshold power at antenna, P_{ma} (dBm)	−88.3	−91.3	−93.3	−93.3
Threshold power at receiver, P_{mr} (dBm)	−86.3	−86.3	−86.3	−86.3

a $\frac{7}{8}$ inner code rate, Δ/T_u of $\frac{1}{8}$ and 16 QAM are assumed, yielding a threshold C/N of 13.9 dB needed to achieve a BER of 2×10^{-4} before R/S decoding. The corresponding payload data rate is 19.35 Mb/s. Since this is just one of many possible scenarios, the entries in this table should not be construed as planning factors. A significant difference between this table and Table 2-1 is the much lower receiver noise figure. In addition, different values of antenna gain and line loss are assumed for the upper and lower portions of the UHF band.

For the ISDB-T system, the theoretical minimum C/N required to achieve a BER of 2×10^{-4} is 16.2 dB, using the same channel bandwidth, modulation, inner code rate, and guard interval ratio[6] as assumed previously for DVB-T. The corresponding payload data rate is 18.93 megabytes per second (MB/s). Thus, in this example the DVB-T system is capable of better performance than the ISDB-T system by about 2.3 dB while achieving a somewhat higher data rate. In fact, the performance difference ranges from 1.4 to 2.7 dB for all possible inner code rates and modulation types. In the hierarchical mode, the DVB-T system requires higher C/N thresholds and achieves lower data rates.

At the time of this writing, an implementation loss of up to 1 dB has been measured on ISDB-T; for DVB-T the measured implementation loss is currently 2.7 dB.[7] As hardware and software developments proceed, performance improvements should be expected. At present, actual performance of both systems is about equal, but the greater potential for improvement is in favor of DVB-T.

EXTERNAL NOISE SOURCES

Although it is standard practice to make calculations as presented in Tables 2-1 and 2-2, this may not tell the complete story. These results represent the minimum power required in an environment limited to random noise, due to the receiver. To obtain the total system noise, the effect of antenna noise temperature, T_a,

[6] "Transmission Performance of ISDB-T," *ITU-R Document 11A/Jyy-E*, May 14, 1999.

[7] Yiyan Wu, "Performance Comparison of ATSC 8-VSB and DVB-T COFDM Transmission Systems for Digital Television Terrestrial Broadcasting," *IEEE Trans. Consumer Electron.*, August 1999.

and the noise contribution of the antenna-to-receiver transmission line must be included. The result is a fictitious temperature that accounts for the total noise at the input to the receiver. When the effects of antenna and line on total are included, the total noise power available at the receiver is

$$N = \frac{kT_aB}{\alpha_r} + (\alpha_r - 1)kT_0B + kT_rB$$

where α_r is the line attenuation factor, T_0 is the ambient temperature, and T_r is the receiver noise temperature. The antenna noise power is attenuated by the transmission line; the noise contribution of the line is added directly to the receiver noise. The receiver noise temperature is related to the noise factor, F, by

$$F = 1 + \frac{T_r}{T_0}$$

Receiver noise factor is related to noise figure by

$$NF = 10 \log F$$

Transmission line loss, L, is related to the attenuation factor by

$$L = 10 \log \alpha_r$$

With the inclusion of these factors, system noise temperature, referenced to the receiver input, is given by

$$T_s = \frac{N}{kB}$$

To illustrate the impact of the external noise sources, the equivalent noise temperature and noise power contributions for each of these components are listed in Table 2-3 for an assumed ambient temperature of 290 K. The receiver noise temperatures are computed from the noise figures given in Table 2-1 for the U.S. DTV system. The sum of all contributions is shown as the receive system noise floor. Two cases are shown. The first is a good approximation for rural areas, based on the curve labeled "rural" in Figure 2-1. The second is based on the curve labeled "suburban." These curves show the increasing effect of impulse noise at the lower frequencies. The antenna noise temperature is assumed to be equal to the values on these curves. The threshold signal required at the input to the receiver under the assumed conditions is also shown in Table 2-3. Since the total system noise already includes the receiver contribution, the threshold receiver signal is determined simply by adding the threshold C/N to the total noise floor.

The results shown for threshold signal level in Table 2-3 are higher than those in Table 2-1 and those normally published in DTV receiver noise budgets. This is because estimates of system noise are often published considering only the receiver noise figure and neglecting the contributions of the external sources through the receive antenna and transmission line-to-system noise.

TABLE 2-3. Antenna, Line, and Receiver Contributions to Noise in U.S. DTV Systems

Component	VHF Low	VHF High	UHF
Case 1: Rural			
Receiver temperature (K)	2610	2610	1450
Line temperature (K)	75	170	440
Antenna temperature (K)	3000	250	24
System noise temperature (K)	5070	2940	1900
System noise floor (dBm)	−93.8	−96.1	−98.0
Minimum receiver power (dBm)	−78.6	−80.9	−82.8
Case 2: Suburban			
Receiver temperature (K)	2610	2610	1450
Line temperature (K)	75	170	440
Antenna temperature (K)	189,000	15,700	1500
System noise temperature (K)	153,000	13,000	2490
System noise floor (dBm)	−79.0	−89.8	−96.9
Minimum receiver power (dBm)	−63.8	−74.6	−81.7

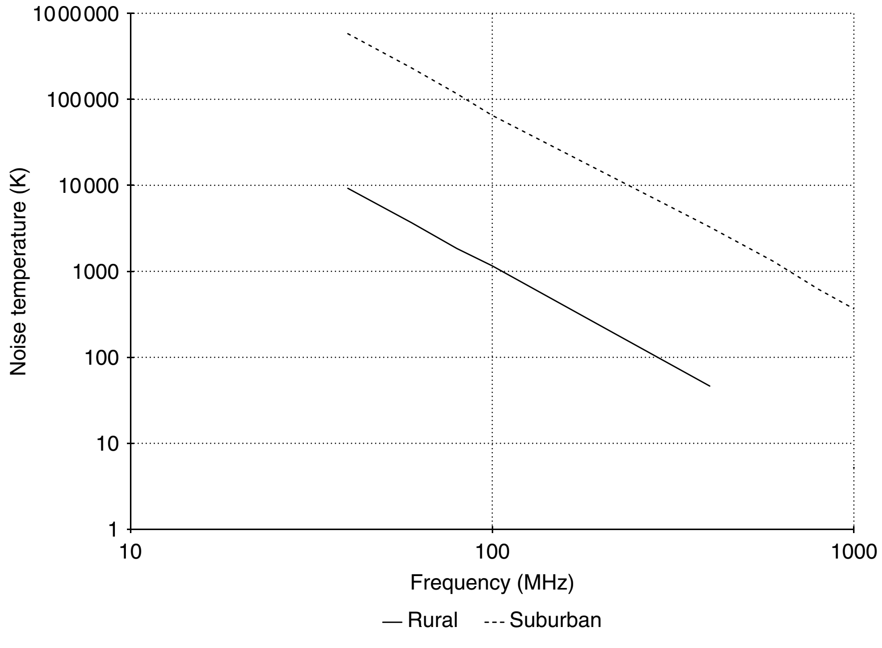


Figure 2-1. External noise temperature. (From *Reference Data for Radio Engineers*, 6th ed., Howard W. Sams, Indianapolis, Ind., 1977, p. 29-2; used with permission.)

Figure 2-1 and the calculations in Tables 2-1 and 2-3 show that the contribution of natural and man-made noise to the antenna and system noise temperature is highly dependent on location, whether in an urban, suburban, or rural environment. In suburban areas the system noise floor may be degraded by external sources by more than 2 dB at UHF; at low-band VHF, the degradation may be over 20 dB. Noise in urban areas may be 16 dB higher than in suburban locations. Rural areas may be quieter than suburban areas by 18 dB or more. Since urban and suburban receivers are more likely to be in areas of high signal strength, there is some justification for using the lowest values for antenna noise temperature to estimate the limits of coverage in many cases. UHF stations may expect to enjoy a 3- to 20-dB noise advantage over low-band VHF stations and a 3- to 6-dB advantage over high-band stations. The advantage due to lower noise level tends to compensate for the higher propagation losses experienced at the higher frequencies.

In practice, the line loss varies with receiver installation as well as frequency. The receiver noise figure varies depending on manufacturer, production tolerances, and frequency. In the tables it is assumed that outside antennas will be used. In those locations where an inside antenna is used, the minimum receive power is increased by the difference in antenna gain. This, too, varies from site to site. The antenna gain varies with manufacturer, production tolerances, and frequency. Thus the threshold receiver power must be understood for what it is — an estimate whose actual value in any given location depends on many site-specific variables.

The higher system noise level due to external sources is qualitatively consistent with field measurement in the United States. In the Charlotte, North Carolina, DTV field tests[8] there were six sites for which no cochannel interference was noted on Channel 6. The average noise floor recorded at these sites was −67.9 dBm; the minimum was −73 dBm and the maximum was −64 dBm. Adjusting these values for the VHF system gain of 25.5 dB results in an average noise floor of −93.4 dBm, a minimum of −98.5 dBm, and a maximum of −89.5 dBm. The equivalent receiver input noise power for the receiver used (NF = 6 dB) was −100.2 dBm, 1.7 dB below the minimum measured value (after adjustment for system gain). The minimum value was evidently measured at a rural location some 21 miles northeast of the transmitter site. Most (but not all) of the locations at which higher noise floors were observed appear to be at more urban or suburban sites. The location at which maximum noise was measured was a part of the Charlotte grid.

For UHF, the average noise floor recorded at the Charlotte field test sites was −71.0 dBm; the minimum was −71.9 dBm and the maximum was −68.2 dBm. Adjusting these values for the UHF system gain of 29.4 dB results in an average noise floor of −100.4 dBm, a minimum of −101.3 dBm, and a maximum of −97.6 dBm. The equivalent receiver input noise power for the receiver used

[8] *Field Test Results of the Grand Alliance HDTV Transmission System*, Association of Maximum Service Television, Inc., September 16, 1994.

(NF = 7 dB) was −99.2 dBm, 2.1 dB above the minimum measured value, 2.4 dB below the maximum measured value, and 1.2 dB above the average measured value (all after adjustment for system gain). From these data it may be concluded that use of only receiver input noise power is a much better predictor of noise floor at UHF. Variation in noise power from location to location is much less at UHF.

The impact of man-made noise at VHF is recognized in the Implementation Guidelines for DVB-T. Noise power is assumed to increase by 6 dB in band I and 1 dB in band III. No allowance is made for man-made noise in bands IV and V.

TRANSMISSION ERRORS

At least three different methods may be used to count transmission errors: segment error rate, bit error rate, and symbol error rate. Symbol error rate is defined as the probability of a symbol error before forward error correction coding. This quantity is often plotted as a function of C/N or the related quantity, E_b/N_0. The relationship between E_b/N_0 and C/N may be derived as follows.

The average carrier power may be written as[9]

$$C = \frac{E_s}{T}$$

where E_s is defined as the energy per symbol and T is the symbol time. The average energy per bit is therefore

$$E_b = \frac{C}{R_b} = \frac{E_s}{TR_b} = \frac{E_s}{\log_2 M}$$

where R_b is the transmission rate in bits per second and M is the number of levels. For example, for 8 VSB, $M = 8$, so that $E_b = E_s/3$. Dividing both sides by N_0, we see that E_b/N_0 is related to C/N by

$$\frac{E_b}{N_0} = \frac{C}{RN_0} = \frac{C}{N}\frac{B}{R_b}$$

Both quantities are usually expressed in decibels, so the latter expression is often written

$$\frac{E_b}{N_0}(\text{dB}) = \frac{C}{N}(\text{dB}) - 10\log\left(\frac{B}{R_b}\right)$$

[9] David R. Smith, *Digital Transmission Systems*, Van Nostrand Reinhold, New York, 1985, pp. 240–241.

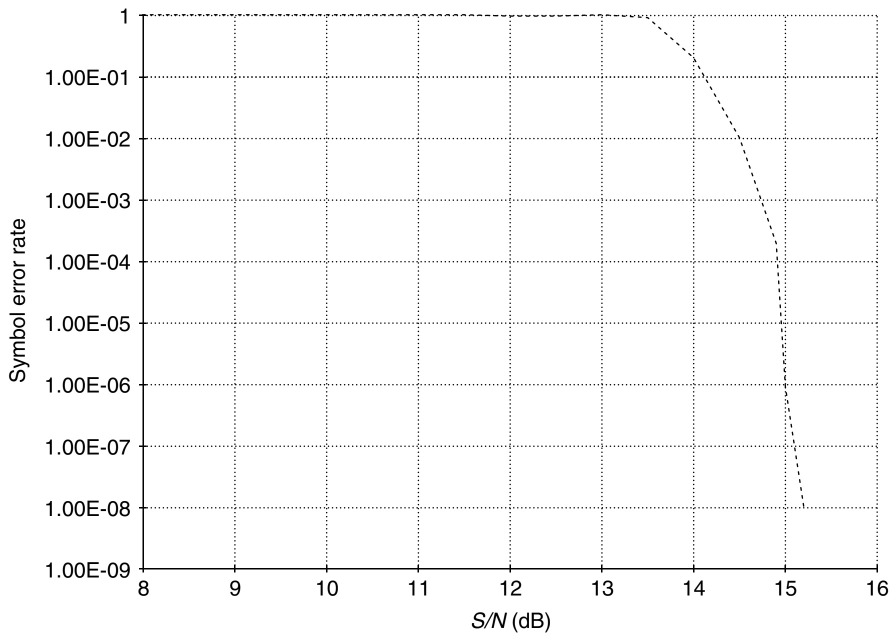


Figure 2-2. Symbol error rate versus S/N. (From Advanced Television Systems Committee, "Guide to the Use of the ATSC Digital Television Standard," *Document A/54*, ATSC, Washington, D.C., Oct. 4, 1995; used with permission.)

The receiver noise bandwidth is assumed equal to the channel bandwidth. This expression allows fair comparison of the relative performance of different systems with differing C/N thresholds and data rates on the basis of E_b/N_0, provided that the error rates are equivalent. If error rates are not equivalent, further adjustment is required. Bit error rate, the probability of a bit error before FEC, is also usually plotted as a function of E_b/N_0.

Segment error rate refers to the probability of an error in a data segment, after forward error correction. Measurements of the segment error rate versus S/N for the 8 VSB terrestrial broadcast mode is shown in Figure 2-2. It is apparent that the system is quite robust until the threshold level is approached. The TOV has been determined to occur for a segment error rate of 1.93×10^{-4}. Recalling that the segment length is 832 symbols and the symbol rate is 10.76 Msymbols/s, it is evident that the TOV corresponds to 2.5 segment errors per second. The equivalent BER is 3×10^{-6} after R/S decoding.

It is instructive to compare the DVB-T and ATSC systems using E_b/N_0. Such a comparison has been made by Wu,[10] who concludes that the ATSC system holds a theoretical advantage over DVB-T of about 1.3 dB. This advantage can be accounted for entirely by the more powerful R/S and convolutional codes

[10] Wu, op. cit., p. 3.

used for the ATSC system. As presently implemented, the advantage is 3.6 dB in the AWGN channel. Measurements on the ATSC system resulted in only a 0.4-dB implementation loss. With improvements, the implementation loss of both systems will be reduced, the DVB-T system having the greater potential improvement.

ERROR VECTOR MAGNITUDE

The quality of the in-band signal may also be expressed in terms of error vector magnitude. This a useful quantitative measure defined as the root-mean-square (RMS) value of the vector magnitude difference between the ideal constellation points, D_i, and the actual constellation points, D_a, of the $I-Q$ diagram, expressed in percent. An error signal vector e_i, may be computed at each symbol time:

$$e_i = D_i - Y_i$$

EVM is usually computed as an average over a large number, N_s, of samples, so that

$$\text{EVM} = \left(\frac{1}{N}\sum_{n=1}^{N_s} |e_i|^2\right)^{1/2} \times 100\%$$

A perfect digital transmission system would exhibit an EVM of 0%.

The inverse relationship between EVM and C/N may be seen by considering the error signal to be noise. The C/N is simply the ratio of the RMS value of the desired constellation points to the RMS value of the noise:

$$\frac{C}{N} = 10\log\frac{\sum_{n=1}^{N_s} D_i^2}{\sum_{n=1}^{N_s} |e_i|^2}$$

The relationship between C/N and EVM is illustrated in Figure 2-3, which is a plot of measured in-band performance at the output of an 8 VSB DTV transmitter. Over most of the range of measurements, EVM is inversely proportional to C/N. Only at low values of EVM is a more complex relationship evident.

Overall, EVM may be considered the best overall measure of in-band DTV performance. It takes into account all impairments that contribute to intersymbol interference (ISI), the underlying cause of symbol and bit errors. ISI is caused by any energy within one symbol time that would interfere with reception in another symbol time. In addition to noise, this energy may be due to dispersion within the channel due to linear distortion or timing errors caused by bandlimiting in the system. The channel response smears and delays the transmitted signal at the receiver. When ISI becomes sufficiently severe, the receiver mistakes the value of the transmitted symbols.

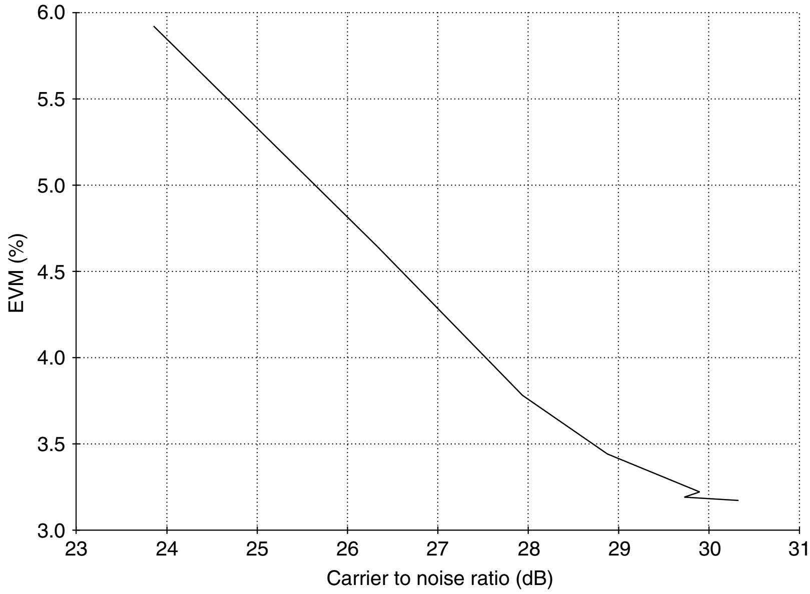


Figure 2-3. EVM versus C/N.

To minimize the effect of dispersion and maximize noise immunity and the resultant ISI in the ATSC system, the square pulses at the input to the 8 VSB modulator are shaped by means of a Nyquist filter. This low-pass linear-phase filter has a flat amplitude response over most of its passband, approximating an ideal low-pass filter. In practice, the ideal low-pass filter with infinitely steep skirts is not physically realizable. Therefore, the response of the Nyquist filter is actually made somewhat more gradual.

The pulse shape at the output of the Nyquist filter is very nearly described by the familiar sinc function,

$$\frac{\sin \pi t/T}{\pi t/T}$$

This function has the property that it is equal to zero at $t = \pm T$, $t = \pm 2T$, $t = \pm 3T$, and so on, but is equal to unity for $t = 0$. Thus pulses occurring at symbol times other than $t = 0$ do not contribute to received symbol power and there is no ISI. The sinc function may be multiplied by any other function without changing the timing of the zeros, thus preserving the property of no ISI. The usual choice is to multiply by a function having a root-raised-cosine response characteristic. The resulting pulse shape is a modified sinc function:[11]

$$\frac{\sin \pi t/T}{\pi t/T} \frac{\cos \alpha_N \pi t/T}{1 - (4\alpha_N^2 t^2/T^2)}$$

[11] Smith, op. cit., p. 210.

This pulse shape also has the property that it is equal to zero at $t = \pm T$, $t = \pm 2T$, $t = \pm 3T$, and so on, but is equal to unity for $t = 0$. As with the sinc function, pulses occurring at symbol times other than $t = 0$ do not contribute to received symbol power and there is no ISI, provided that the half-power bandwidth is $1/2T$. The tails of this pulse decrease at a faster rate than the ideal low-pass filter, so that timing precision is not as critical.

The half-power bandwidth of this filter is often referred to as the Nyquist bandwidth, which for the U.S. DTV system is 5.38 MHz. The full filter bandwidth is greater by a factor of $1 + \alpha_N$, so that the channel bandwidth is 1.1152×5.38 or 6 MHz. Outside this frequency band, the response is zero.

EYE PATTERN

The eye pattern is a convenient method for visually and qualitatively assessing the ISI and C/N performance of a digital transmission system. The signal is displayed on an oscilloscope set to trigger at the symbol time. The persistence of the scope creates a composite of all possible waveforms. At each level of the signal, the overlapping waveforms produce a pattern that resembles the human eye. The degree to which the eyes are open is a measure of the ISI and hence signal quality. The eye openings should be greatest at the sampling time. Eyes open 100% correspond to an EVM of 0%.

Approximate quantitative measure of C/N may be made from visual assessment of the eye pattern.[12] The signal amplitude is represented by the center-to-center distance between symbol levels, V. Noise is represented by the accumulated thickness of the intersecting lines at each symbol level, ΔV. The log ratio of these distances, $20 \log(V/\Delta V)$, is an estimate of C/N ratio. For example, if $\Delta V = 0.1$ V, the C/N value is 20 dB. Similar closing of the eyes in the horizontal dimension may be an indication of timing jitter.

INTERFERENCE

Although the noise floor is a useful concept for estimating the maximum extent of coverage, in the real world interference is often present, tending to place further limits on coverage. Interfering signals may originate with cochannel and adjacent channel stations. Signals further removed in frequency may be either harmonics or intermodulation products (IPs). In any case, the level of these signals at their source is usually outside the control of the stations with which they interfere. During the transition period, the interference may come from both analog and digital TV signals. Only the effects of interference on the digital television signal will be discussed.

[12] Luobin and K. Cassidy, "Analyze QAM Signals with Oscilloscope Eye Diagram," *Microwaves and RF*, January 1998, p. 115.

The geographical locations, channels, and effective radiated power (ERP) of existing analog stations are presently fixed. Digital stations will generally be located at or near analog sites with sufficient power to replicate analog service. These factors and the channel assignments of the digital services are the starting point for interference analysis. Propagation of interfering signals is dependent on the same factors that affect the desired signal. Given the AERP, tower-siting parameters, and operating channel, the time-varying signal level at a specific location depends on distance, topographical factors, atmospheric conditions, and sources of multipath. The methods described in Chapter 8 may be used to estimate these levels using the parameters of the interfering station(s).

COCHANNEL INTERFERENCE

Cochannel signals are the desired signals for stations in other markets. The digital TV receiver detects a cochannel digital signal as just another source of noise. When the C/N value is less than threshold due to the combination of thermal sources and interference, reception will fail. The thermal noise and interfering signal powers are additive, so that the $C/(N+I)$ threshold is increased in inverse proportion to the cochannel interference. This is illustrated in Figure 2-4, which shows the C/N threshold as a function of carrier-to-interference ratio, C/I, for a pair of 8 VSB DTV stations operating on the same channel, assuming an

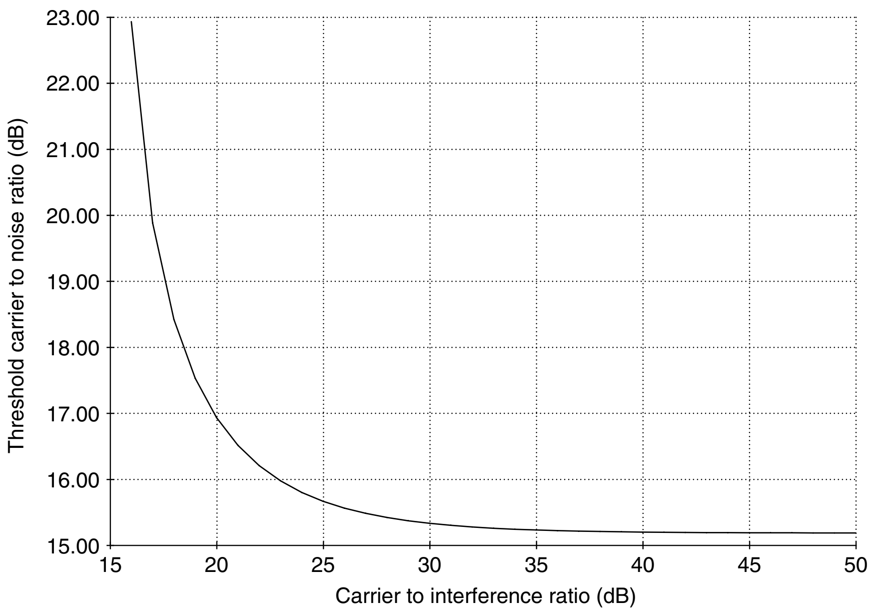


Figure 2-4. Threshold C/N versus C/I.

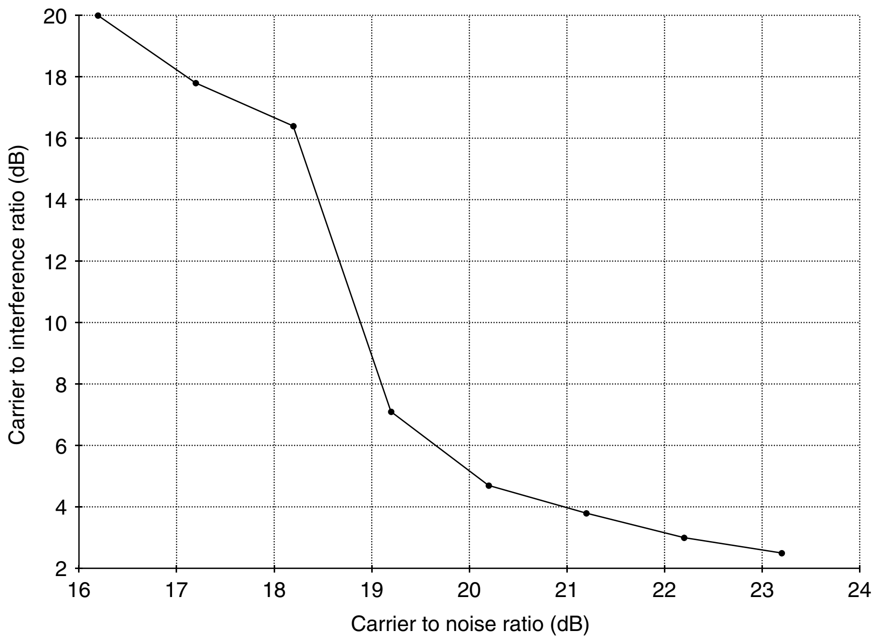


Figure 2-5. C/I versus C/N. (From *DTV Express Training Manual*; used with permission.)

omnidirectional receiving antenna. When the C/I value is high, say greater than 35 dB, the threshold C/N approaches 15.2 dB, the interference-free value. As the interference increases, the threshold C/N increases correspondingly. It is apparent that near the edge of noise-limited coverage, the minimum C/I is much higher than in areas where C/N is high. At the edge of noise-limited coverage, no interference can be tolerated.

During the transition period, cochannel analog stations may also affect digital TV service. Results of tests at the Advanced Television Test Center (ATTC)[13] on the 8 VSB system using a noncommercial laboratory DTV demodulator are shown in Figure 2-5. From these tests it is apparent that for low DTV carrier-to-noise ratios ($C/N < 19.2$ dB) the critical desired-to-undesired ratio, D/U, is 7 dB. That is, the desired signal must exceed the undesired signal by 7 dB for acceptable DTV reception. At the critical D/U value the DTV picture was impaired with black squares frozen in time and the audio was muted. At lower D/U, the picture, sound, and data failed completely. The laboratory DTV demodulator was equipped with a NTSC cochannel rejection filter, which automatically switched in when the C/I value fell below 16.5 dB. At high carrier-to-noise ratios ($C/N > 23$ dB), the critical D/U is 2 dB. At the time of this writing, no data are available with consumer-grade receivers. A

[13] Charles Rhodes, *DTV Express Training Manual*, Harris Corporation, Melbourne, FL, p. 2–9.

NTSC cochannel rejection filter may not necessarily be present in all commercial receivers.

The referenced test data were the basis of channel allocations in the United States, where a cochannel analog-to-digital protection ratio of 1.8 dB is used. For DVB-T, the corresponding protection ratio is 4 dB. In general, no improvement should be expected from the use of precise carrier offset by analog transmitters since interfering signals may come from any one of many stations. Some analog stations may offset the visual carrier by 10 kHz, with a tolerance of up to ± 1 kHz.

In the absence of a NTSC offset, offsetting the DTV pilot frequency by 28.615 kHz above its normal frequency has the effect of placing the NTSC visual, chroma, and aural signals near the nulls of the NTSC reject filter in DTV receivers[14] equipped with these filters.

ADJACENT CHANNEL INTERFERENCE

An adjacent channel signal may be the desired signal for another station or the result of third- or higher-order intermodulation products generated in the power amplifier of other transmitters. Whatever the source, these signals appear as spurious sidebands in the adjacent channel just outside the desired channel. When generated by digital TV transmitters, these components appear as noise. The FCC and DVB-T masks define strict limits on this interference at the output of the transmitter system. The noise remaining after application of the mask adds to the noise from other sources.

There is very little difference between the effect of interference from upper or lower digital TV sidebands. The effect is much the same as that of digital-to-digital cochannel interference, except that the levels are offset by 45 dB. This is approximately equivalent to the total noise power relative to the average in-band power resulting from use of an emissions mask.

ANALOG TO DIGITAL TV

For analog transmitters on adjacent channels, the major concern is for the visual, color, and aural carriers interfering with the digital TV signal via the adjacent channel reject bands of the receiver. There are no artifacts produced until a critical D/U value is reached. The critical D/U is -48 dB for a lower adjacent NTSC and -49 dB for an upper adjacent NTSC signal. Thus, the undesired signal may be as much as 49 dB greater than the desired signal. At a D/U above this level, reception fails abruptly — there is no picture, sound, or data.

The sidebands of an adjacent analog station may also interfere with the in-band digital TV signal. For example, the specification for lower sideband reinsertion

[14] C. Eilers and G. Sgrignoli, "Digital Television Transmission Parameters — Analysis and Discussion," *IEEE Trans. Broadcasting*, Vol. 45, No. 4, December 1999, pp. 365–385.

for NTSC transmitters is only 20 dB. To the writer's knowledge, no tests have been conducted to measure the magnitude of this effect.

TRANSMITTER REQUIREMENTS

To provide the desired C/N at the receiver, it is necessary that the transmitter system produce a signal free of noise and of both linear and nonlinear distortions at sufficiently high power to cover the service area. The average power is the parameter to which the performance of a digital TV transmitter system is referenced. A typical time-domain signal envelope for an 8-VSB modulated signal is shown in Figure 2-6. Note the great variability in the envelope peaks. Obviously, there is no regularly recurring peak corresponding to the familiar sync pulse of analog television. Neither is there a predictable peak envelope signal level. However, the average power level is constant. A similar envelope could be plotted for OFDM signals. Therefore, a digital TV station's power is normally stated in terms of the average transmitter output power (TPO) or average effective radiated power.

Aside from being a constant for any combination of video, audio, and data, there are many advantages to using average power to characterize a digital television transmission system. Techniques for measurement of average power are well developed. For high-power systems, use of a calorimeter provides a useful means of calibrating average system output power. Other average reading instruments using methods that depend on development of a dc level due to heating of the sensing element may also be used. Because the average power level is constant in time and the spectrum is constant over most of the channel bandwidth, relative power levels may readily be determined by integration of the radiated spectrum.

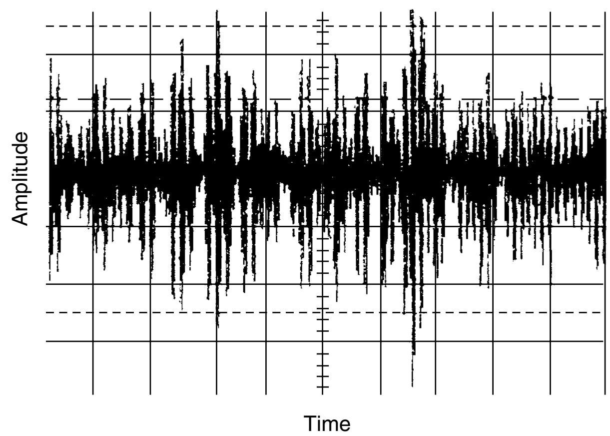


Figure 2-6. Time-domain envelope of RF DTV signal. (Photographed by Bob Plonka, Harris Corporation; used with permission.)

Although average power is used to establish transmitter and station ratings, knowledge of the peak power is not unimportant. Sufficient headroom must be allowed in the intermediate and final amplifiers of the transmitter as well as any other nonlinear components in the transmission chain to avoid excessive compression — thus the requirement for very linear power amplification. Moreover, the peak power rating of transmission lines, filters, RF systems, and antennas must be specified properly to avoid voltage breakdown.

The model numbers of most transmitters are usually assignal in terms of a peak rating. This seems to be due more to the tradition of rating analog transmitters in terms of their peak sync rating than of any definite measurement. Nevertheless, measurement of peak power is much more difficult and in many cases not necessary. It is common practice to estimate the peak-to-average ratio (PAR) based on reasonable assumptions, and to multiply this factor by the average power to obtain an estimate of peak power. PAR may be estimated on the basis of the cumulative distribution function (CDF). This is illustrated using data from 8 VSB tests in Figure 2-7. Note that PAR does not exceed 7.5 dB about 99.99% of the time. For the COFDM signal, the corresponding PAR is about 10 dB. These figures are considered to be adequate estimates for specifying the allowable compression in a power amplifier. Peak-to-average ratios in 8 VSB transmitters of up to 11 dB have been reported. The higher values might be used to select a transmission line conservatively based on the breakdown voltage or

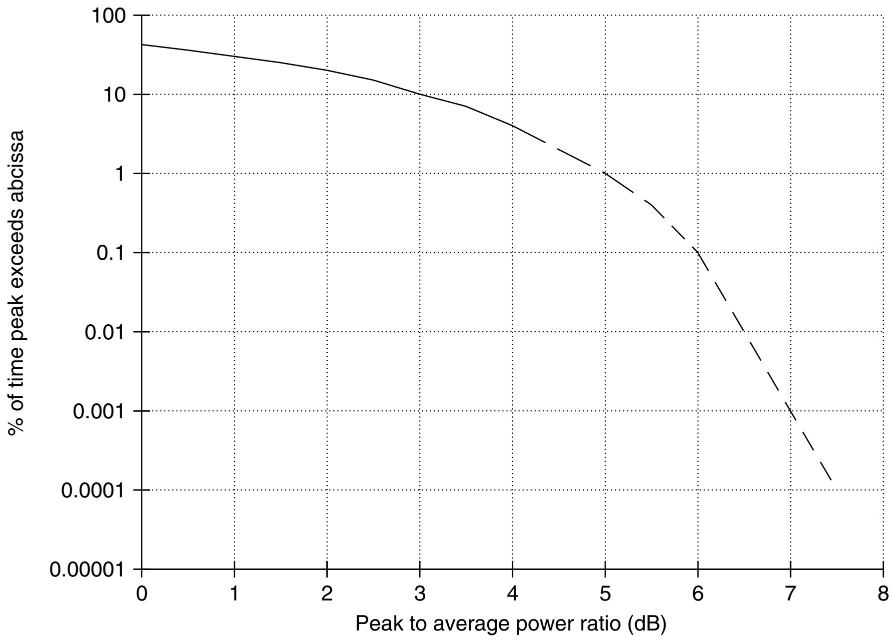


Figure 2-7. Typical CDF of 8 VSB signal. (From *DTV Express Training Manual*; used with permission.)

to specify the headroom in components that must be extremely linear, such as the exciter.

The large signal peaks relative to the average signal level result in compression in the output stages of practical power amplifiers. The resulting nonlinearity produces higher-order intermodulation products, which are observed as spectral spread. Figure 2-8 shows the spectrum of an 8 VSB signal measured at the output of a typical power amplifier. The signal within the channel bandwidth is essentially constant except for a spike in the spectrum on the left side. This spike is due to the presence of the pilot signal. Outside the channel bandwidth there is evidence of spectral regrowth. In an ideal linear amplifier, the energy in this region would be limited to the levels generated in the exciter. The measured level in a practical amplifier depends on the extent to which the amplifier is driven into compression during the signal peaks as well as the specific nonlinear characteristic of the amplifier.

Spectral spread is an extremely important parameter, in that it determines the level of interference to digital and analog stations allocated to adjacent channels. From the viewpoint of the adjacent channel, the out-of-band energy simply becomes another source of noise. Even if there is no adjacent channel allocation, the FCC and DVB-T specifications require stringent limits to out-of-band signals. For example, the mandated FCC radiation mask is shown in Figure 2-9. The 0-dB reference is set at the average in-band power level. In practical terms the mask requires that all signals outside the 6-MHz channel allocation be 36.7 dB below the average in-channel level, decreasing in linear fashion to 99.7 dB below the in-channel level at frequencies 6 MHz above and below

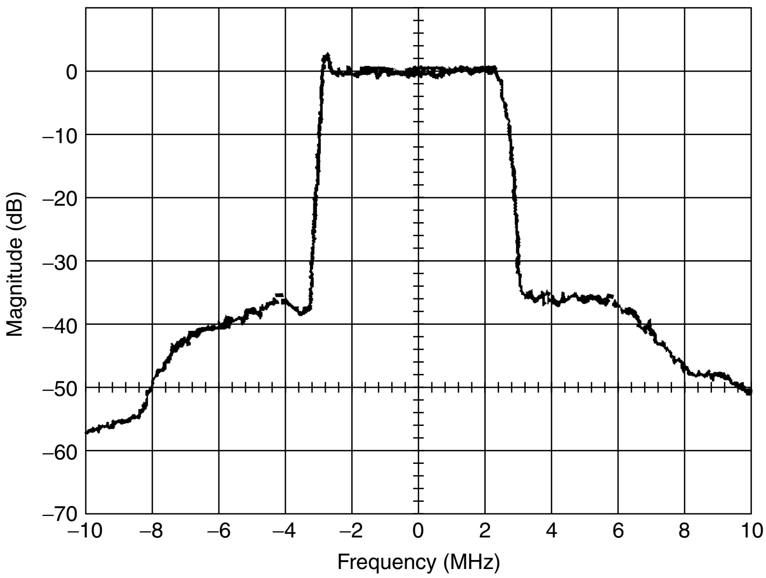


Figure 2-8. Typical transmitter output spectrum.

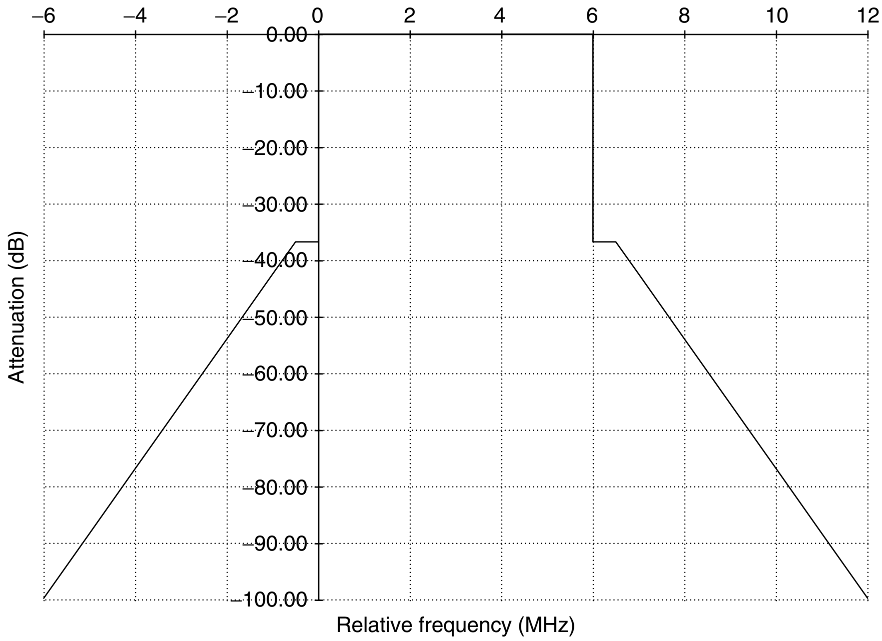


Figure 2-9. FCC emissions mask.

the channel edge. As the mask is plotted here, it is assumed that the resolution bandwidth of the spectrum analyzer is fixed to a sufficiently small value (say, 30 kHz) to properly present the in-band and out-of-band measurements. Similar masks for the DVB-T system are shown in Figures 3-10 and 3-11. Needless to say, to achieve the levels required by the applicable masks requires a power amplifier with adequate headroom, precise power control, stable and precisely controlled precorrection circuits, and a well-designed output bandpass filter.

AM-to-AM conversion is the primary mechanism by which spectral regrowth occurs. This term is used to describe the degree to which the transmitter output voltage is directly proportional to the input. Output phase may also be a function of input level. The deviation from ideal linear phase is described as AM-to-PM conversion. A perfectly linear transmitter would produce no AM/AM or AM/PM.

Consider a practical amplifier (or any other quasilinear component). In the time domain, the output, S_o, of a third-order nonlinearity may be described as a function of the input, S_i, as follows:

$$S_o = gS_i + g_3S_i^3$$

where g is the gain of the amplifier in the linear region of the transfer function and g_3 represents the degree of third-order nonlinearity. Thus, if g_3 is zero, the amplifier is ideal, having no AM/AM or AM/PM conversion. If g_3 is nonzero, the amplifier is nonlinear. The resulting AM/AM and AM/PM will give rise to third-order intermodulation products and spectral regrowth.

The nonlinear transfer function is illustrated in Figure 2-10 for the case of $g = 10$ (20 dB) and $g_3 = -1$. Over most of the input range, the output increases in direct proportion to the input. Note that when the input signal is about unity (0 dB), the compression is approximately 1 dB. As the input signal increases to a value of 1.8 (5 dB), no further increase in output occurs. The compression at this point is about 3.4 dB. Beyond this point, increases in the input produce less output signal. Some practical amplifiers, such as klystrons, actually exhibit this type of nonlinear characteristic, in which the output is reduced when the input signal increases beyond a limiting value.

AM-to-PM conversion may be illustrated by considering the third-order term in the transfer function to be complex. In this case we simply write

$$g_3 = g_{3I} + jg_{3Q}$$

where g_{3I} is the in-phase component of the nonlinear term and g_{3Q} is the quadrature component. Nonlinear phase and amplitude are illustrated in Figure 2-11 for the case of $g_{3I} = g_{3Q} = -1$. As the input signal increases, the output phase lags. At the 1-dB compression point, the incidental phase shift is about $-7°$, typical of many practical amplifiers.

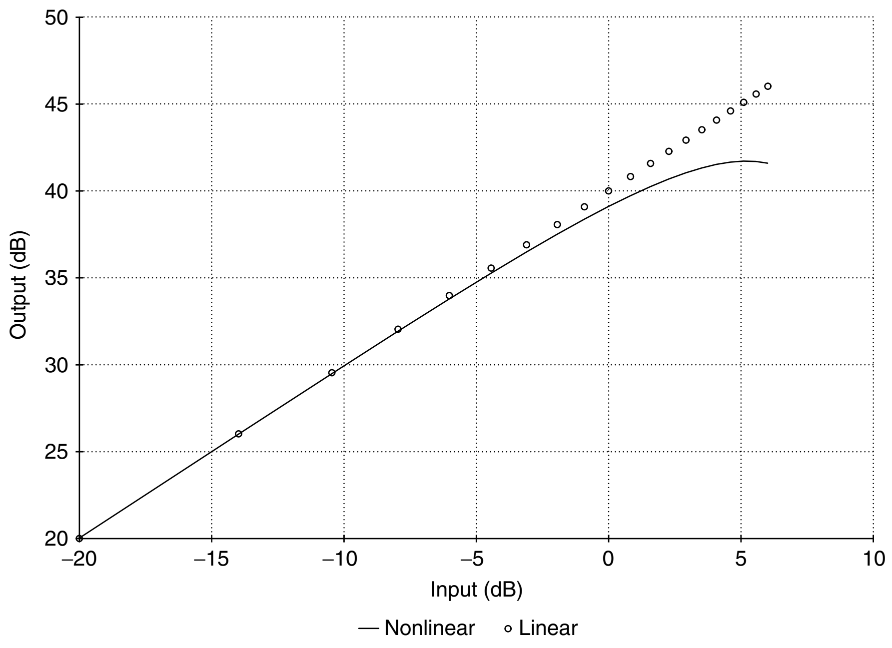


Figure 2-10. Nonlinear amplification.

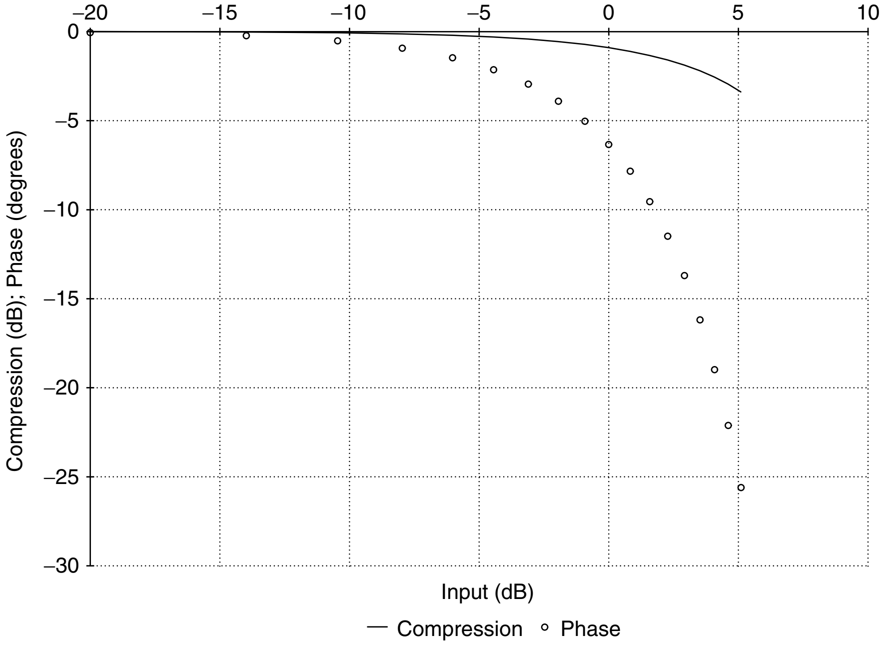


Figure 2-11. Nonlinear phase.

The relationship between AM/AM and spectral regrowth may be illustrated by considering the input signal, S_i, to be composed of two tones, S_1 and S_2, at angular frequencies ω_1 and ω_2, respectively; that is,

$$S_i = S_1 + S_2$$

When these signals are inserted into the expression for S_o, the result is

$$S_o = g(S_1 + S_2) + g_3(S_1 + S_2)^3$$

Expanding the cubic term and after some algebra, we obtain

$$(S_1 + S_2)^3 = S_1^3 + S_2^3 + 3S_1^2 S_2 + 3S_1 S_2^2$$

It is apparent that the fundamental signals have been preserved and amplified. However, additional signals, S_1^3, S_2^3, $3S_1^2S_2$, and $3S_1S_2^2$, have been generated. It is well known that these new signals are, among others, at frequencies $3\omega_1, 3\omega_2$, $2\omega_1 - \omega_2$, and $2\omega_2 {-} \omega_1$. The third harmonic signals at $3\omega_1$ and $3\omega_2$ are well outside the channel and may be filtered. However, the difference signals at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ are in and near the channel and represent intermodulation products (IPs) or spectral regrowth. Recognizing that the digital signal may be described as a continuous spectrum, it is apparent that the continuous spectral regrowth due to nonlinearity shown in the measured data is to be expected. In fact, the amplitude of the intermodulation products could be computed using this type of model.

In the absence of other noise sources, spectral regrowth places a minimum value on the transmitted in-band noise floor. Consider the case when ω_1 and ω_2 are closely spaced and near the center of the channel, say at 2.5 and 3.5 MHz above the channel edge. In this case the IPs will be 1 MHz above and below the fundamental tones, well within the channel bandwidth. These intermodulation products represent noise with respect to the desired signal. Again considering the digital signal as a continuous spectrum, it is evident that transmitter system nonlinearity places an upper limit on the carrier-to-noise ratio.

There are noise sources other than thermal and third-order products due to nonlinear distortions within a transmitter system. For example, quantization noise in digital-to-analog converters and oscillator phase noise contribute to total noise power. In a well-designed and well-maintained transmitter system, these sources should be small.