NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Electrical and Electronic Engineering

AM BROADCASTING

Graduation Project
EE-400

Student: Alamgir Hasan Riaz (970814)

Supervisor: Prof. Dr.
Fakhreddin Mamedov

Lefkoşa- 2001

ACKNOWLEDGEMENT

I would like to express my gratitude and thanks to all my teachers that played important part in my educational life. I would like to specially thank to my supervisor Professor Dr. Fakhreddin Mamedov who supported me in preparing this project.

ABSTRACT

Amplitude modulation is a simple, efficient method for transmitting information. The original idea for creating a radio signal goes back to James Clerk-Maxwell, an English physicist who theorized the existence of electro-magnetic energy in 1873. Amplitude modulation operates on a specific frequency known as a carrier. This signal never changes in power. The operating power of a broadcasting station is the carrier power. The radio signal is generated in the form of a sine wave. Figure below, left depicts its wavelength and amplitude characteristics. The number of wavelengths occurring in one second is the wave's frequency, measured in cycles per second, or Hertz.

In communications, we often need to send information, referred to as signal, from one point to another. For example, the information can be voice, music or video, as in radio and television broadcasting. To achieve this, at the point of origin, the signal is multiplied by a sinusoidal waveform referred to as the carrier. This process is called modulation. At the point of reception, the signal is extracted from the modulated carrier, a process we call demodulation.

The introduction of stereophonic transmission to AM broadcasting has allowed it to become more competitive with FM stereo broadcasting. However, just as the AM stereo transmission process is very much different than that of FM stereo, so is the requirement for proper audio processing of AM stereo. Special processing requirements are needed in order to maximize both good monaural compatibility and high quality stereophonic transmission simultaneously. Digital transmission methods offer interesting advantages, especially in frequency ranges which, until today, have been used differently.

TABLE OF CONTENTS

AM BROADCASTING

CHAPTER I

1.1 Introduction

Amplitude Modulation is the Modulation in which the amplitude of a carrier wave is varied in accordance with some characteristic of the modulating signal. Amplitude modulation implies the modulation of a coherent carrier wave by mixing it in a nonlinear device with the modulating signal to produce discrete upper and lower sidebands, which are the sum and difference frequencies of the carrier and signal. The envelope of the resultant modulated wave is an analog of the modulating signal. The instantaneous value of the resultant modulated wave is the vector sum of the corresponding instantaneous values of the carrier wave, upper sideband, and lower sideband. Recovery of the modulating signal may be by direct detection or by heterodyning.

This thesis will be in six parts. The introduction part will follow with introduction to the modulation and demodulation in general. The second part will be about the theory of amplitude modulation. This will be followed by part three which is about modulation and demodulation. Part four will be about the application of AM. The chapter 5 will be about a proposal for digital transmission with AM band. The theses will end with conclusion.

First lets be acquainted with AM. Amplitude Modulation (AM) is a broadcast system that seems to have been neglected over the past two decades.

In recent years it has been tagged "Ancient Modulation," and criticized for being full of noise. Many critics have defined it as "low fidelity" when comparing it to FM, and there is some truth to that.

Yet AM has a lot of things going for it. There are more than 4,700 AM radio stations in the United States. The frequency band in which our AM service is located produces a significant groundwave that permits the transmission of reliable signals over a vast area with relatively little power compared to what must be supplied to an FM signal operating in the VHF band. The AM signal travels up and down over hills - obstacles that give VHF signals trouble.

If a new form of digital radio is successful, the AM broadcaster, operating at a frequency with these attributes, has much in its favor.

The medium-wave frequencies on which AM operates also reflect radio signals radiated skyward back to earth. This often causes interference at night. In the early days of broadcasting, the FCC, and the Federal Radio Commission before it, regulated night power and antenna design to limit interference. Night "skip" was something many people actually listened for, attempting to hear programs hundreds of miles away. Some still do, but with so many syndicated talk programs on the AM band at night, the listener often finds the same programming already available in his or her community. Sporting events still have people tuning in to distant stations at night. Clear-channel stations still provide the nighttime service, but it is far less important than it was a half-century ago.

Radio broadcasting has changed drastically over the past few years. Individuals who entered this career just five or 10 years ago find it to be much different than the labor-intensive business it once was. Management no longer has to make sure someone is at the studio and transmitter site to keep program continuity. Advancements in computer technology allow an operator to keep a radio station operating for days unattended.

While this trend may limit careers for budding DJs and talk-show hosts, it does put more of demand on management.

Today's sophisticated broadcast systems require attention, and often they are neglected until the dreaded "dead-air" occurs. Usually, panic ensues, at which point managers usually start calling on anyone who has any knowledge of the system.

Many stations are not prepared to substitute locally originated material. The announcers are not there, and neither is the programming, be it music or talk. The staff managing today's radio stations must know something about how the audio and radio transmission system operates.

New studio and transmitter equipment using computer technology often is beyond the scope of the old "workbench repair" that took place in the past. Often, it is often not in the interest of the broadcaster to purchase the test equipment necessary to make these repairs; often it is not in management's interest to hire an engineer who understands all these concepts and to keep this person up to date by paying for training when necessary. Engineers don't like to hear that, but in this age it is true. It is easier to call the manufacturer and get advice than try to troubleshoot on your own.

Normally a transmitter will run, unattended, for long periods of time without trouble. New components today have incredibly long lives. However, when no engineer is on site, management should inspect the transmitter and antenna site periodically. Management should check how the studio system is functioning as well. An individual need not have experience in electronics to detect problems. A person with a good ear and an ability to read meters correctly is an important asset to make sure the station is running without problems. Such an employee may be able to detect problems before they become a serious threat to station operation.

CHAPTER II: THEORY OF AM

Amplitude modulation is a simple, efficient method for transmitting information. The original idea for creating a radio signal goes back to James Clerk-Maxwell, an English physicist who theorized the existence of electro-magnetic energy in 1873. His theories were proven by Heinrich Hertz, who actually generated and received radio waves in his laboratory in 1888. Hertz did not follow up his work with any practical applications.

The radio wave is the product of an electric current flowing through an unterminated wire or antenna. As the alternating current flows to the end of the wire and then back, a magnetic field is created perpendicular to the current flow. If the wire is spread apart as illustrated in Figure 2.1, the magnetic field will radiate away from the transmission line.

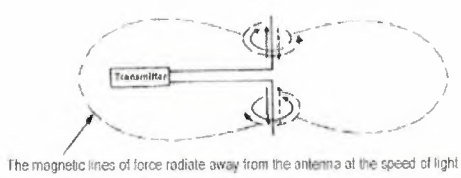


Figure:2.1

Spreading the wires apart creates an antenna. The radio wave is moving away from the antenna at the speed of light. Originally this magnetic field would be intercepted by the receiver antenna, and through electro-magnetic induction, produce a small current that would actuate a relay creating a clicking sound. Morse Code was sent this way.

But some people wanted to do more than just transmit code. One could say that the frequency that was transmitted with code was the carrier. To transmit voice, more than the carrier had to be sent.

with code was the carrier. To transmit voice, more than the carrier had to be sent.

It was discovered that if audio signals were converted to electric current variations, through a microphone, these signals could be added to the radio frequency and decoded in a receiver. Initially these audio signals were capacitively or inductively coupled to the radio frequency. Pioneers like Reginald Fessenden and Lee DeForest demonstrated audio transmissions around the United States and Europe.

However, it was another individual, more closely associated with FM, who improved amplitude modulation and made it practical for broadcasting. Edwin Howard Armstrong invented the principle of regeneration or oscillation. This allowed the alternator to be retired. The totally electronic transmitter was at hand. Armstrong also invented the superheterodyne receiver, making radio reception simple and reliable.

2.1 Amplitude-modulated Transmitters

Figure 2.2(a) below shows the block diagram of a typical AM transmitter. The carrier source is a crystal-controlled oscillator at the carrier frequency or a submultiples of it. This is followed by a tuned buffer amplifier and a tuned driver, and if necessary frequency multiplication is provided in one or more of these stages.

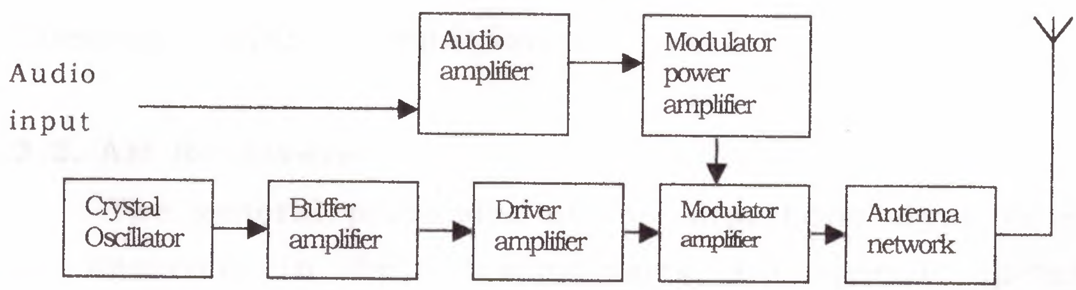
The modulator circuit. used is generally a class C power amplifier that is collector modulated as described in above Section. The audio signal is amplified by a chain of low-level audio amplifiers and a power amplifier. Since this amplifier is controlling the power being delivered to the final RF amplifier, it must have' a power driving capability that is one-half the maximum power the collector supply must deliver to the RF amplifier under 100% modulation conditions. A transformer-coupled class B push-pull amplifier is usually used for this purpose.

Low-power transmitters with output powers up to 1 kW or so may be transistorized, but as a rule the higher-power transmitters use vacuum tubes in the final amplifier stage, even though the low-level stages may be transistorized. In some cases where the reliability and high overall efficiency of the transistor are mandatory, higher powers can be obtained by using. several lower-power transistorized amplifiers in parallel. The system is complicated, and usually the vacuum-tube version will do the same job at lower capital cost,

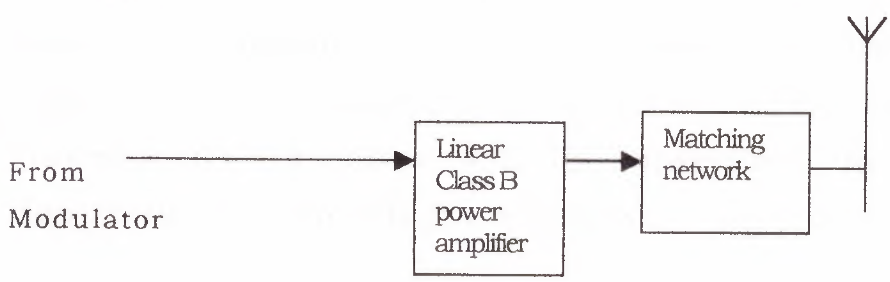
Sometimes the modulation function is done in one of the low-level stages. This allows low-power modulation and audio amplifiers, but it complicates the RF final amplifier. Class C amplifiers cannot be used to amplify an already modulated (AM) carrier, because the transfer function of the class G amplifier is not linear. The result of using a class C amplifier would be an unacceptable distortion of the modulation envelope. A linear power amplifier, such as the push-pull class B amplifier, must be used to overcome this problem Figure 2.2(b). Unfortunately, the efficiency of this type of amplifier is lower than that of the comparable class C amplifier, resulting in more costly equipment. Larger tubes or transistors must be used that are capable of dissipating the additional heat generated.

The output of the final amplifier is passed through an impedance matching network that includes the tank circuit of the final amplifier. The R of this circuit must be low enough so that all the sidebands of the signal are passed without amplitude/frequency distortion, but at the same time must present an appreciable attenuation at the second harmonic of the carrier frequency. The bandwidth required in most cases is a standard 3 dB at ±5 kHz around the carrier. For amplitude-modulation broadcast transmitters, this response may be broadened so that the sidebands will be down less than 1 dB at 5 kHz where music programs are being broadcast and very low distortion levels are desired, or special sharp-cutoff filters may be used. Because of the high power levels present the output, this is not usually an attractive solution.

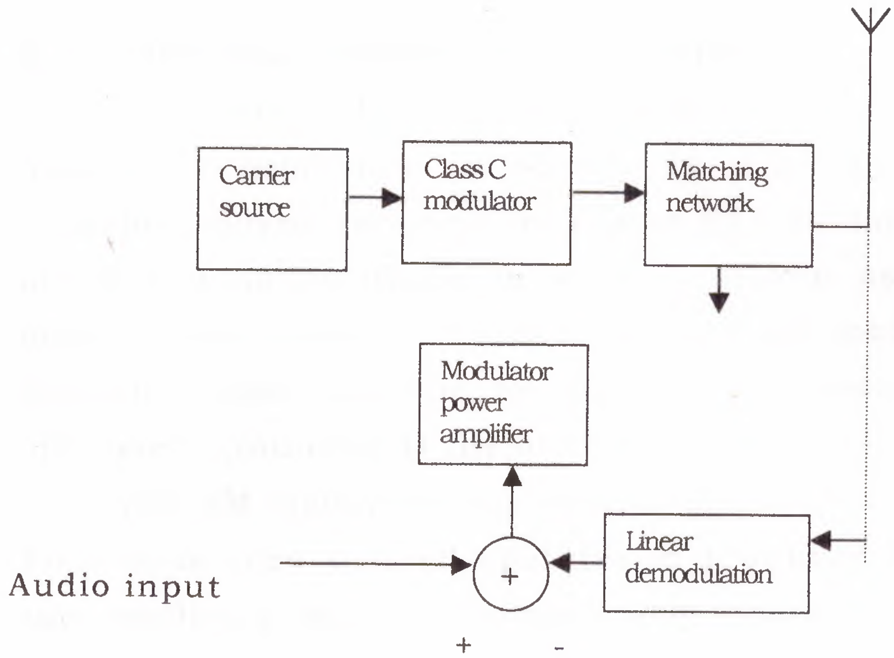
Negative feedback is quite often used to reduce distortion in a class C modulator system. The feedback is accomplished in the manner shown in Fig. 2.2 (c), where a sample of the RF signal sent to the antenna is extracted and demodulated to produce the feedback signal. The demodulator is designed to be as linear in its response as possible and to feed back an audio signal that is proportional to the modulation envelope. The negative feedback loop functions to reduce the distortion in the modulation.



(a)



(b)



(c)

Figure: 2.2

Amplitude modulated transmitters: (a) transmitter with a modulated class C final power amplifier; (b) linear class B push-pull power amplifier used when modulation takes place in a low-level stage; (c) negative feedback applied to linearize a class C modulator.

2.2. AM Receivers

The general principles of the superheterodyne receiver are described in the following parts, and specific operating details of the AM envelope detector are also discussed in chapter IV. Most receivers in use today are assembled from discrete components, although there is a trend toward the use of integrated circuits for subsections in the receiver. Therefore, in this section, a very commonly encountered transistorized receiver will. be described, followed by the description of two integrated circuit-type receivers.

2.2.1 Discrete Component AM Receiver

The circuit for a standard broadcast receiver using discrete components is shown in Fig. 2.3. This is a superheterodyne receiver, transistor Ql functioning as both a mixer and an oscillator in what is known as an autodyne mixer. The oscillator feedback is through mutual inductive coupling from collector to emitter, the base of Q1 being effectively grounded at the oscillator frequency.

The AM signal is coupled into the base of Ql via coil Ll. Thus it is seen that Ql operates in grounded base mode for the oscillator while simultaneously operating in grounded emitter mode for the signal input.

Tuned IF transformer Tl couples the IF output from Q1 to the first 1F amplifier Q2. The output from Q2 is also tuned-transformer-coupled through T2 to the second IF amplifier Q3. The output from Q3, at IF, is tuned-transformer-coupled to the envelope detector D2, which has an RC load consisting of a 0.01-q, F capacitor in parallel with a 25-kS2 potentiometer. This potentiometer is the manual gain control, the output from which is fed to the audio preamplifier Q4. The audio power output stage consists of the push-pull pair Q5, Q6.

Automatic gain control (AGC) is also obtained from the diode detector D2, the AGC filter network being the 15-kS2 resistors and the 10-WF capacitor (Fig. 2.3). The AGC bias is fed to the Q2 base.

Diode D1 provides auxiliary AGC action. At low signal levels, D1 is reverse biased, the circuit being arranged such that the collector of Ql is more positive than the colleetor of Q2. As the signal level increases, the normal AGC bias to Q2 reduces Q2 collector current, resulting in an in- crease in Q2 collector voltage. A point is reached where this forward biases

D1, the conduction of D1. then damping the Tl primary and so reducing the mixer gain.

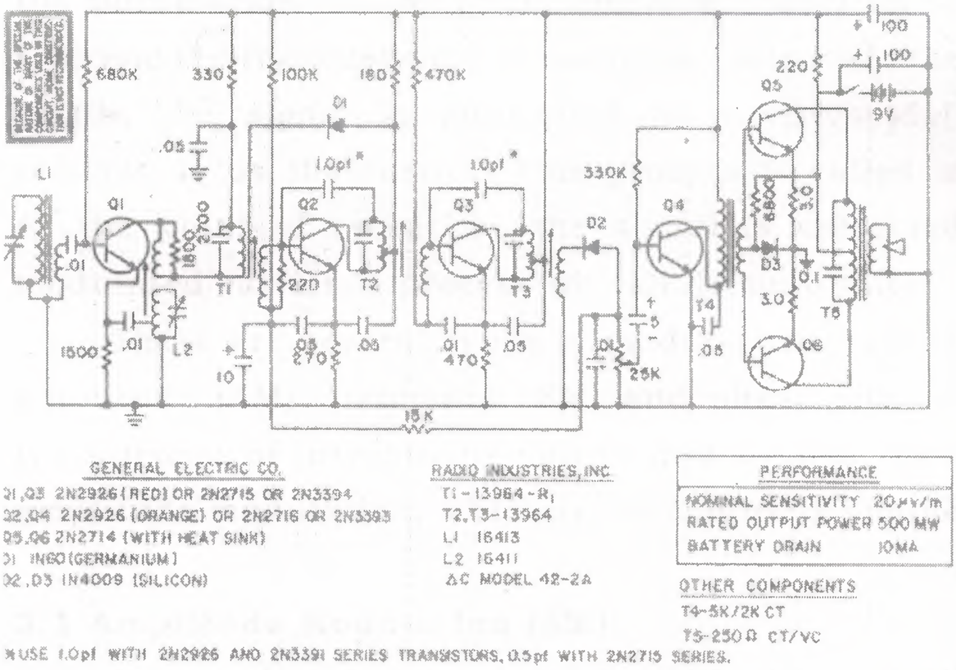


Figure: 2.3

CHAPTER III: MODULATION AND DEMODULATION

In communications, we often need to send information, referred to as signal, from one point to another. For example, the information can be voice, music or video, as in radio and television broadcasting. To achieve this, at the point of origin, the signal is multiplied by a sinusoidal waveform referred to as the carrier. This process is called modulation. At the point of reception, the signal is extracted from the modulated carrier, a process we call demodulation.

There are several types of modulation. Among them are, amplitude (AM), frequency (FM) and phase modulations. For the purpose of introducing you to this subject, we will choose amplitude modulation, which is conceptually simpler.

3.1 Amplitude Modulation (AM):

Let us denote our signal by x(t), which may be either periodic or non-periodic. We indicate the carrier by: $c(t) = \cos(\omega_c t+\phi_c)$
The modulated waveform will be:

$$y(t) = x(t)c(t) = x(t)\cos(\omega_c t+\phi_c)$$

This is represented by figure 3.1:

$$g(t) = y(t)c(t) = x(t)\cos^2(\omega_c t+\phi_c)$$

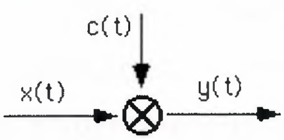


Figure: 3.1

At the destination, the received signal, y(t), is again multiplied by the same carrier sinusoid with identical phase.

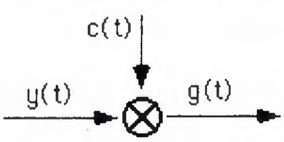


Figure: 3.2

Using the trigonometric identity:

$$\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$$

we obtain:

$$g(t) = \frac{1}{2}x(t) + \frac{1}{2}x(t)\cos(2\omega_c t + 2\phi_c)$$

The function g(t) has two parts. The first is one half the signal we intended to transmit. The second part is the product of this signal with a sinusoid having a frequency twice the carrier frequency. If the carrier frequency is much higher than any frequency contained in the signal, x(t), it is easy to separate these two pieces. This is achieved by passing g(t) through a low-pass filter, shown in block diagram of figure 3.3.

Figure: 3.3

The low-pass filter allows frequencies below a cut-off value pass through and blocks higher frequencies. A simple low-pass filter can be made from a resistor and a capacitor as shown in the following figure 3.4.

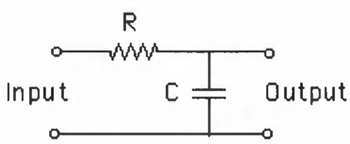


Figure 3.4

The capacitor has the property that it offer very little resistance to high frequency signals, essentially appearing as a short circuit to them so they do not pass through the filter. On the other hand, it appears as an open circuit to low-frequency signals and allows them to pass through.

If we define $H(\omega)$ as the ratio of the amplitude of the output sinusoid to the input sinusoid at frequency ω, it is given by the following equation:

$$H(\omega) = \frac{1}{\sqrt{1+\left(\frac{\omega}{\omega_0}\right)^2}}$$

where,

$\omega_0 = \frac{1}{RC}$

This is the critical frequency that divides the high and low frequency range for a particular low-pass. At this frequency, the amplitude of the output sinusoid reduces to 0.707 the value at zero frequency (1 in this case). The figure 3.5 shows H(ω).

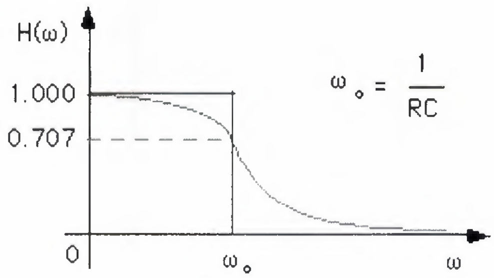
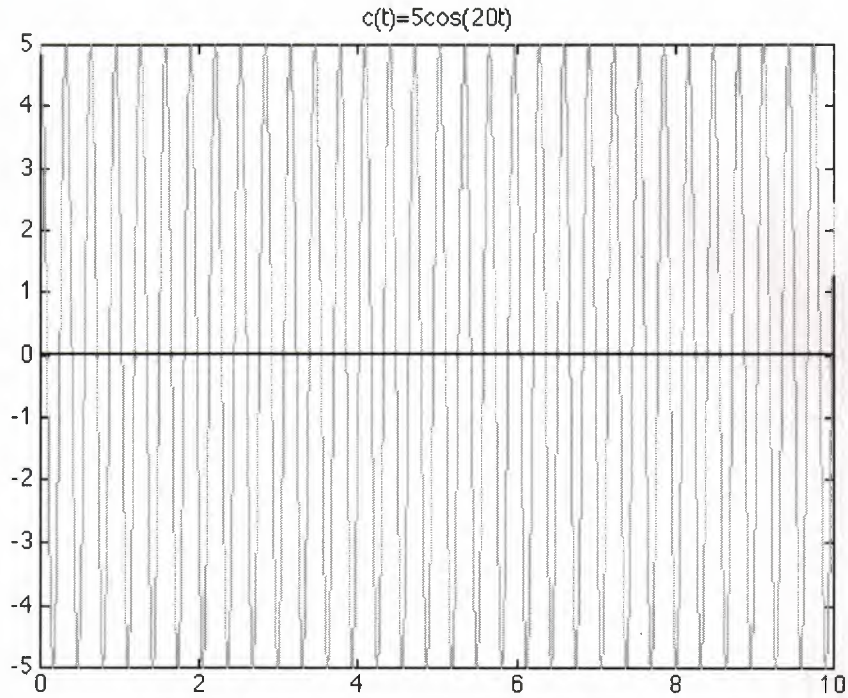


Figure: 3.5

The curve H for an ideal low-pass filter that passes all frequencies bellow ω_0, without any reduction of their amplitude, and completely blocks all frequencies above this value.

Example: In this example we choose the carrier as c(t) = 5cos(20t) and the signal as x(t) = cos(2t).

Carrier Frequency



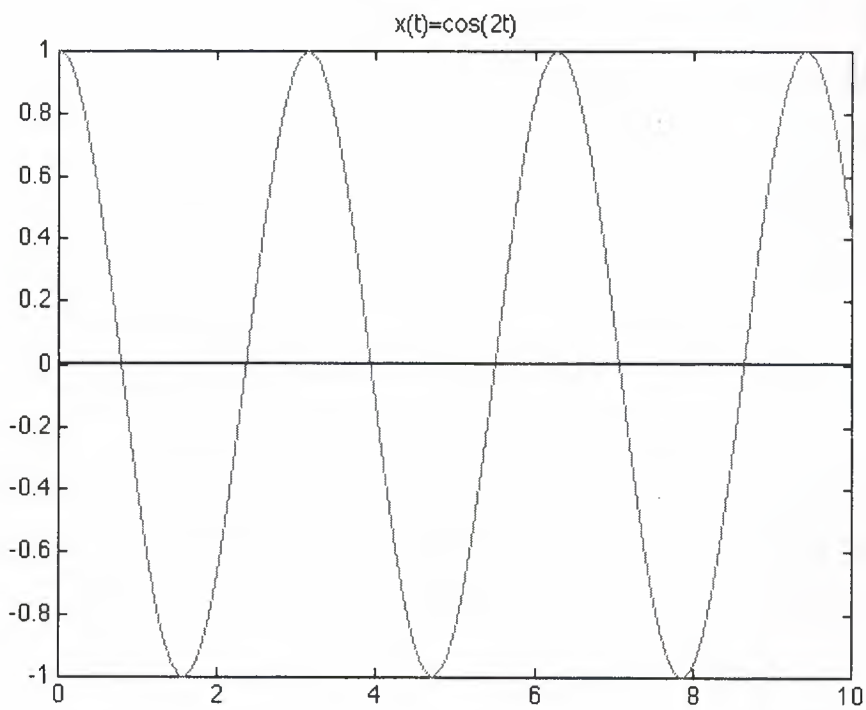


Figure 3.6

The modulated signal, $y(t) = x(t)\ c(t)$ is shown in figure 3.7.

AM Modulated Signal

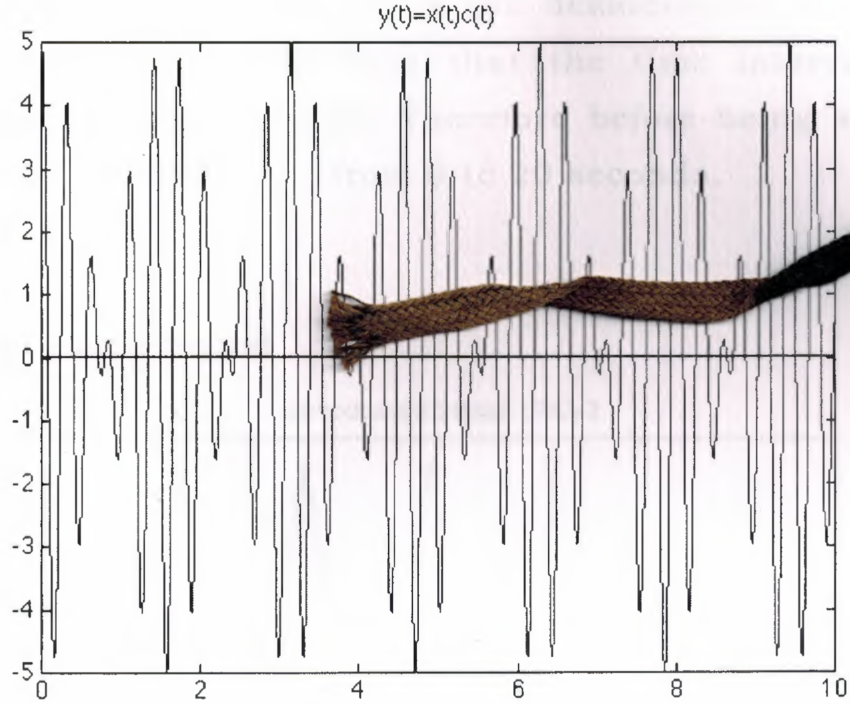


Figure: 3.7

To demodulate the received signal, we must pass it through a low-pass filter with its cut-off frequency higher than 2 radian/s (signal frequency) and lower than 40 radian/s (twice the carrier frequency). We can accomplish this using MATLAB in the following manner:

1. Create the function,

$h(t) = e^{-\omega_0 t}$

where ω_o is the cut-off frequency we choose for the low-pass filter.

2. Create a function g(t) = c(t) y(t).

3. Enter the following command in MATLAB:

```
>> xd = conv(h , g)
```

The resulting function x_d is the demodulated signal and is shown in figure 3.8. Note that the time interval for this function is from 0 to 20. Therefore before being able to plot we must re-define "t" , from 0 to 20 seconds.

Demodulated Signal

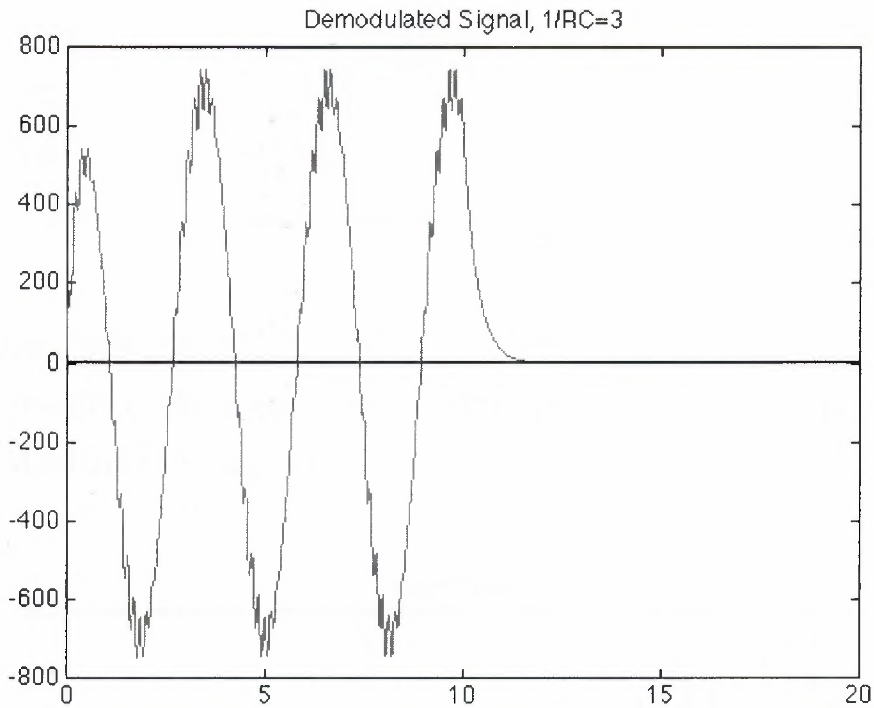


Figure: 3.8

Example:

In this example we choose out signal to be a ramp, x(t) = 0.01t and keep the same carrier as in example I.

Signal

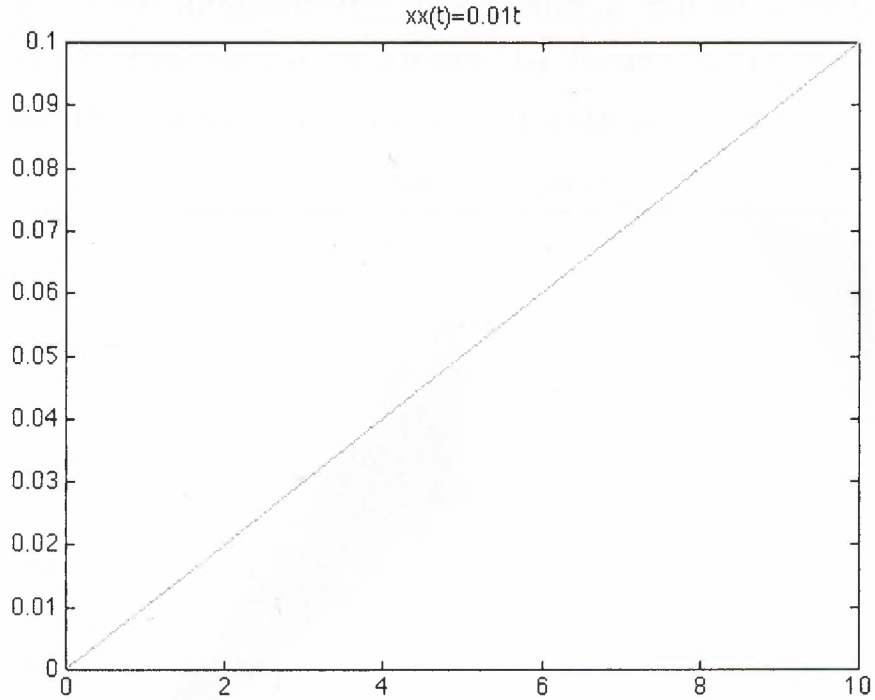
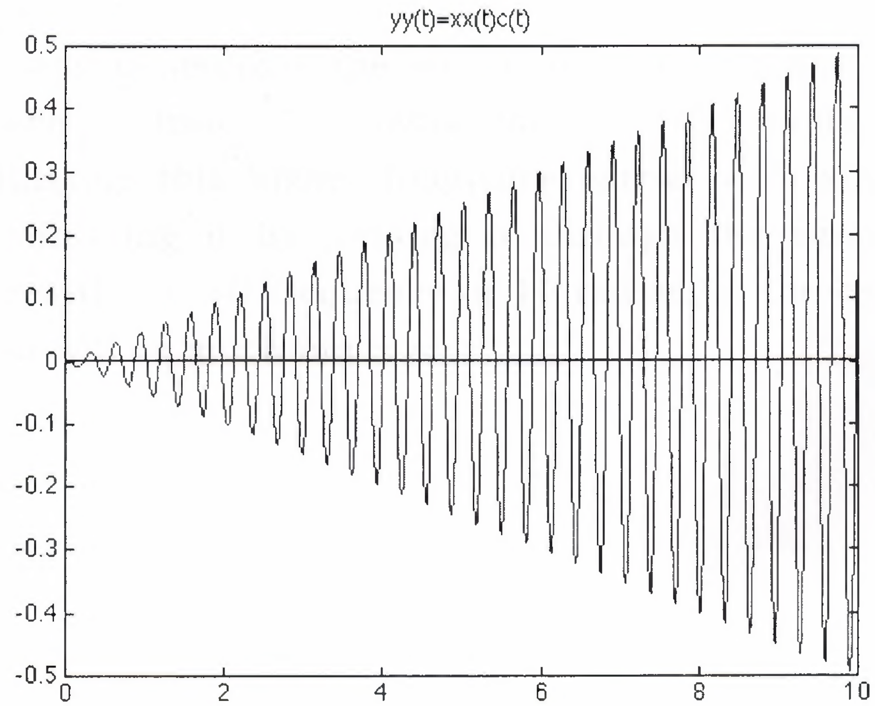


Figure: 3.9

The modulated signal y(t) = x(t) c(t) is shown in figure 3.9 : AM Modulated Signal



We will now demodulate this using a cut-off frequency of 10 radian/s. The result is shown in figure 3.10. As we can see the result of demodulation is not very satisfactory.

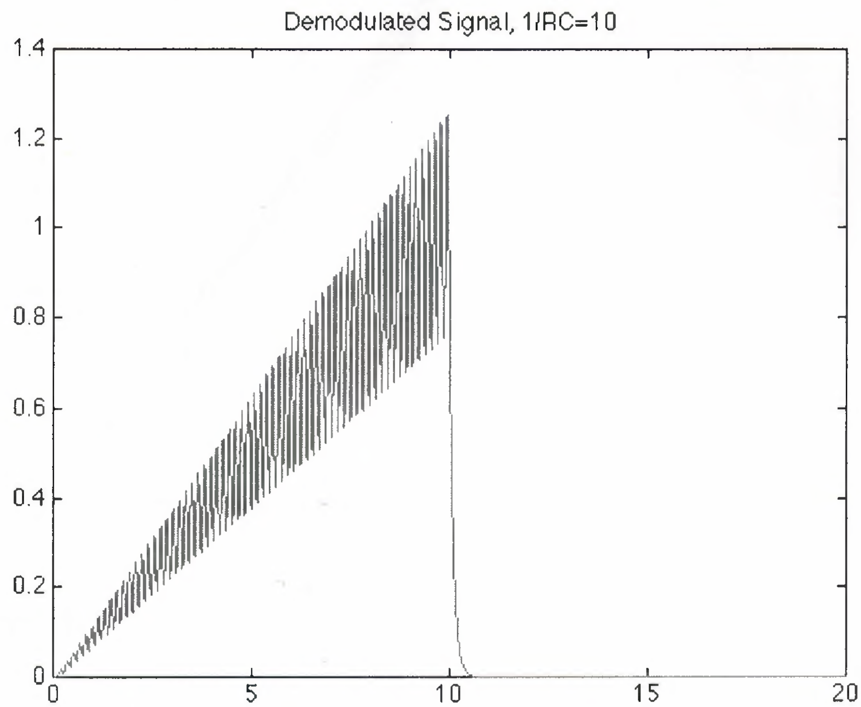


Figure 3.10

One way to improve the situation is to increase the carrier frequency from 20 radian/s to 100 radian/s. After modulating this higher frequency carrier with x(t) and then demodulating it by passing it through the same low-pass filter with cut-off frequency of 10 radian/s, the demodulated signal will be much improved:

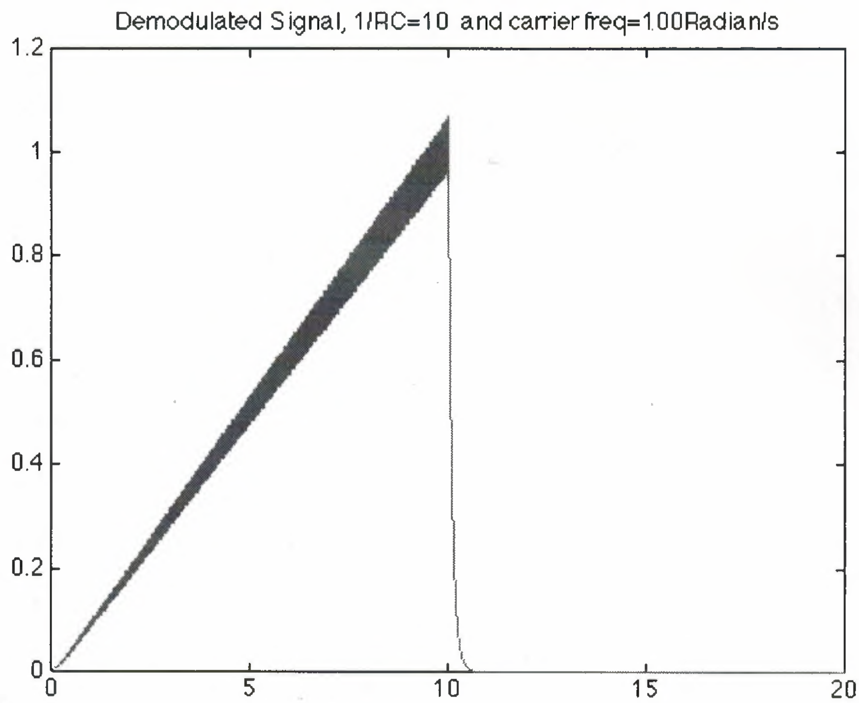


Figure: 3.11

Example:

This time our signal is an exponential of the form:

$x(t) = e^{-t}$

Signal

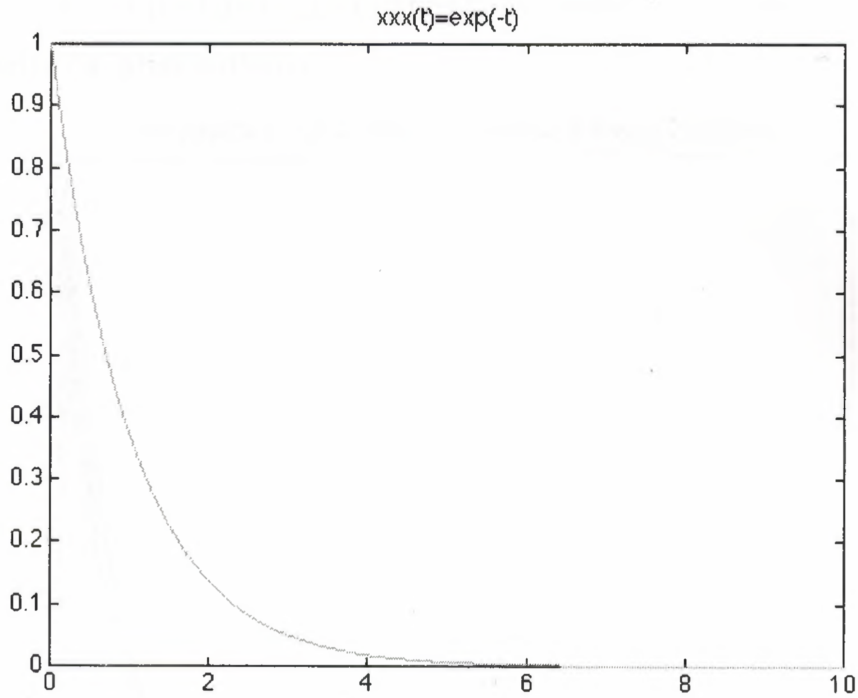
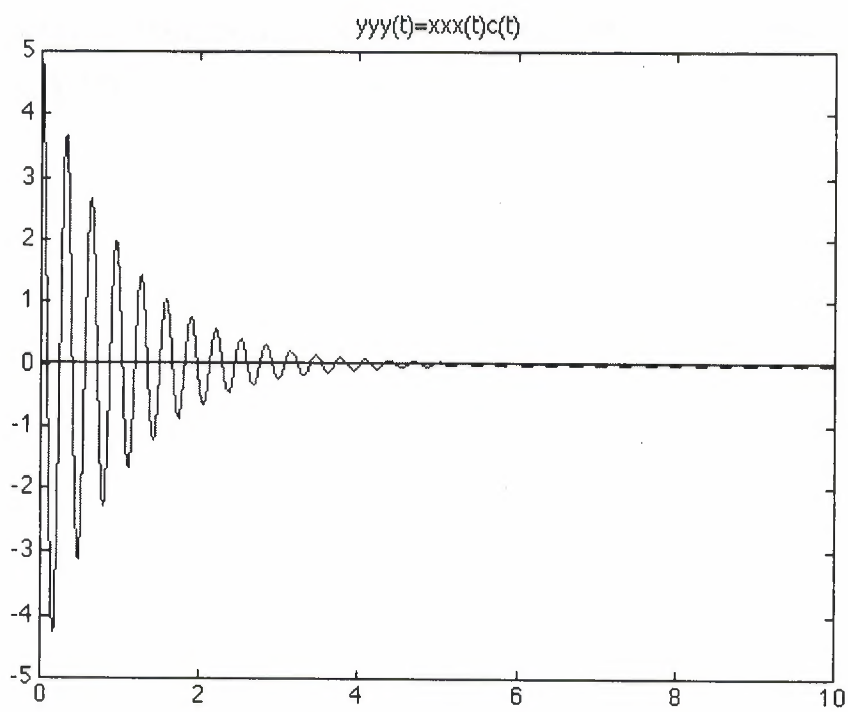


Figure: 3.12

We will keep the carrier the same as in example I. The modulated signal is shown in figure 3.13 here:

AM Modulated Signal

Figure:3.13



Again, we will first demodulate it with the filter cut-off at 10 radians/s and obtain:

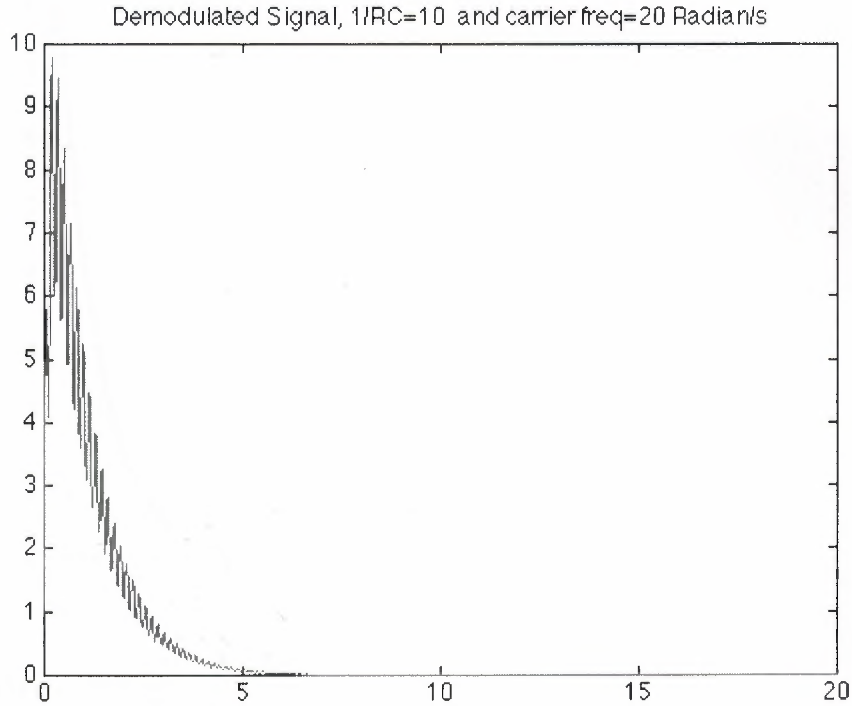


Figure 3.14

Once again we will try to improve the process by increasing the carrier frequency to 100 radian/s. The result is shown in figure 3.15:

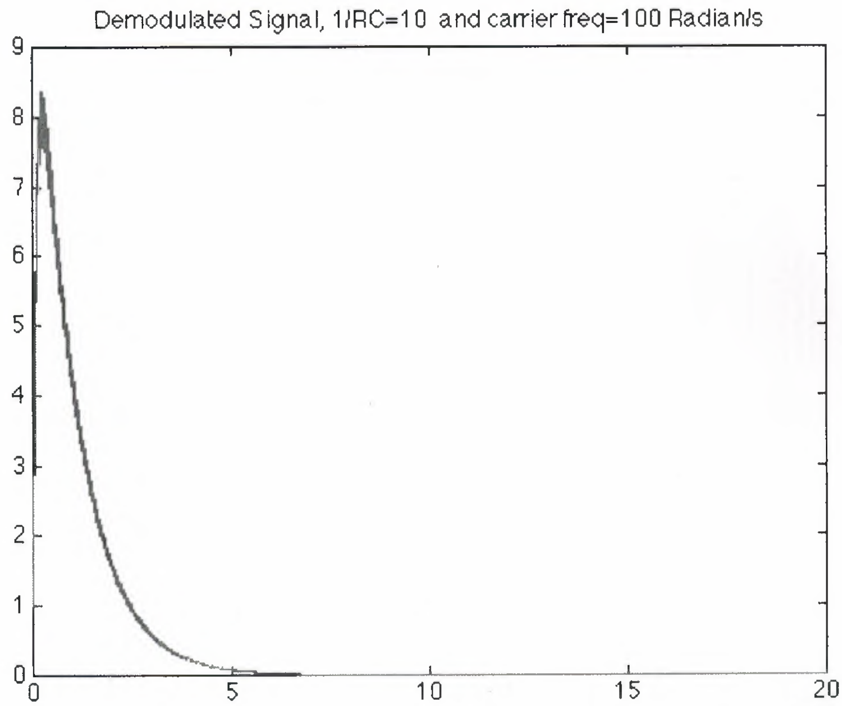


Figure: 3.15

B. Frequency Modulation (FM):

In this form of modulation, the signal x(t) is superimposed on the frequency of the carrier. Therefore, the carrier frequency becomes dependent on time in one of the following two ways:

$\omega_c(t) = \omega_c + x(t)$

or,

$\omega_c(t) = \omega_c \cdot x(t)$

Example:

Let us take the carrier to be c(t) = 5cos(30t) and the signal to be constant, i.e., x(t) = 0.1 . Then: $\omega_c(t) = \omega_c \cdot x(t) = 30 \times 0.1 = 3$ Rad/s

We will show the carrier, the signal and the modulated signal in figure 3.16. The demodulation in the case of FM is to complicated for this course, so we will not attempt to deal with it here.

Carrier Frequency

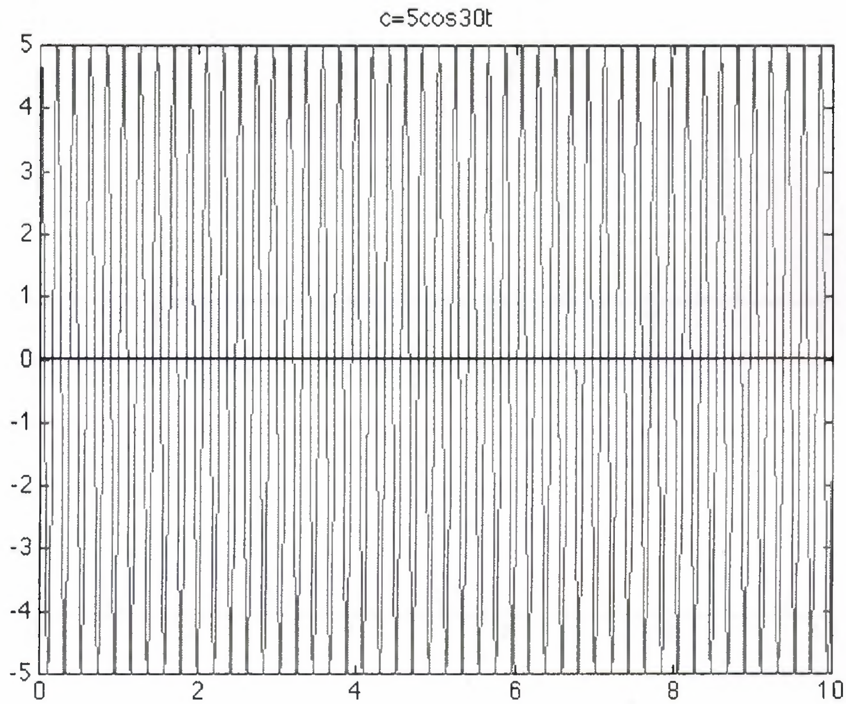


Figure: 3.16

Signal

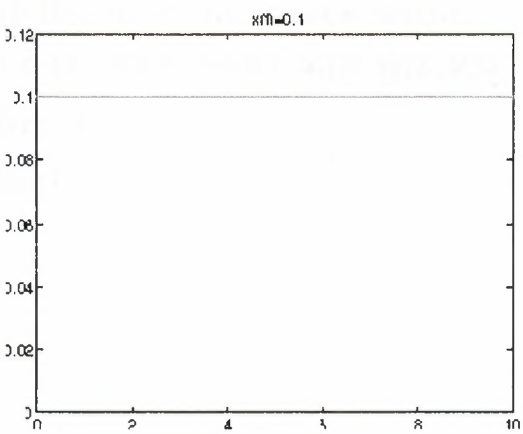


Figure: 3.17

FM Modulated Signal

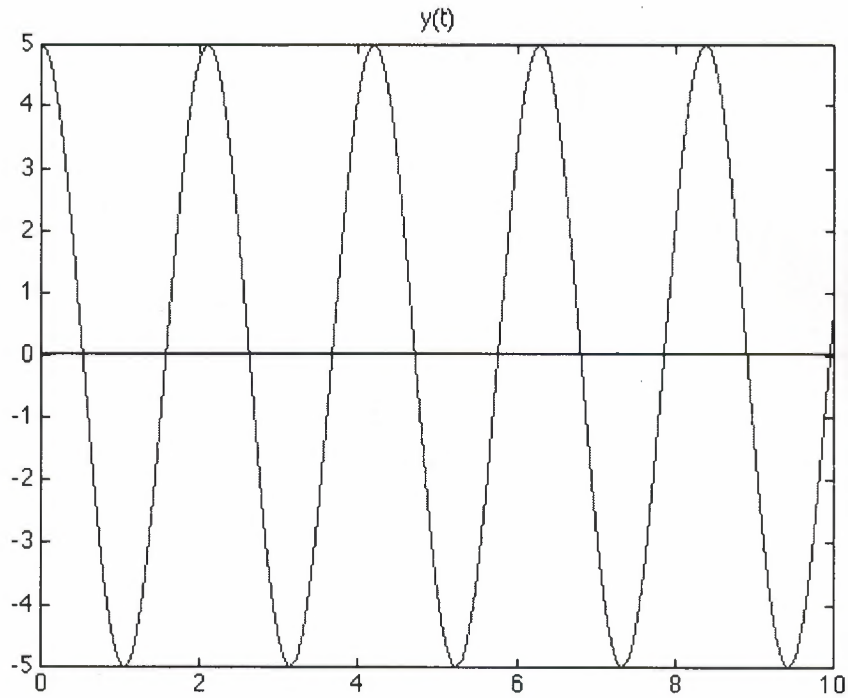


Figure: 3.18

Example:

In this example we keep the same carrier but choose a sinal that linearly increases with time, i.e., a ramp of the form, x(t) = 0.01t. The sinal and the FM modulated signal are shown in figure 3.19:

Signal

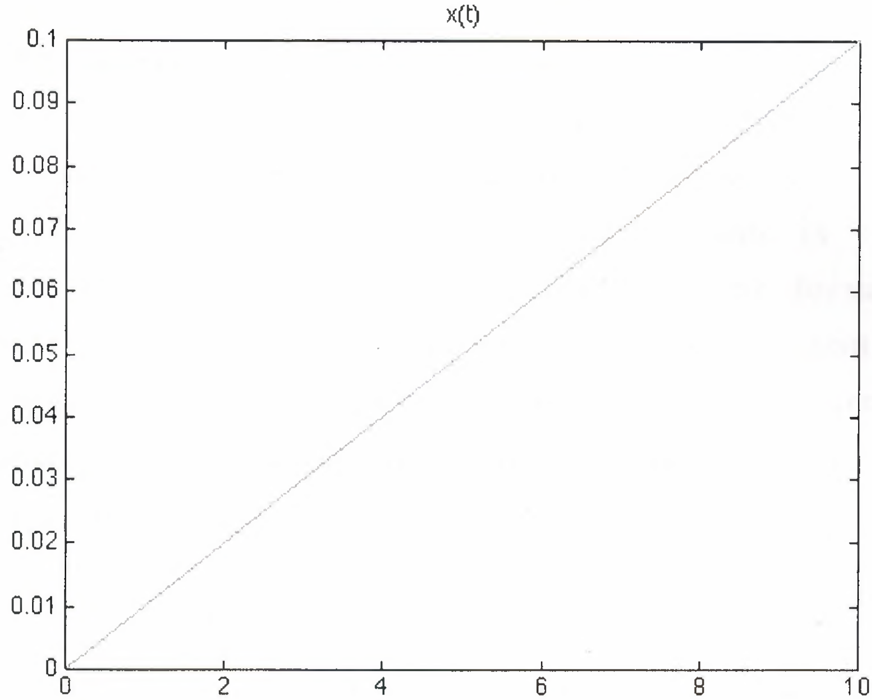


Figure: 3.19

FM Modulated Signal

$\omega_c(t) = \omega_c \cdot x(t) = 30 \times 0.01t = 0.3t$ Rad/s

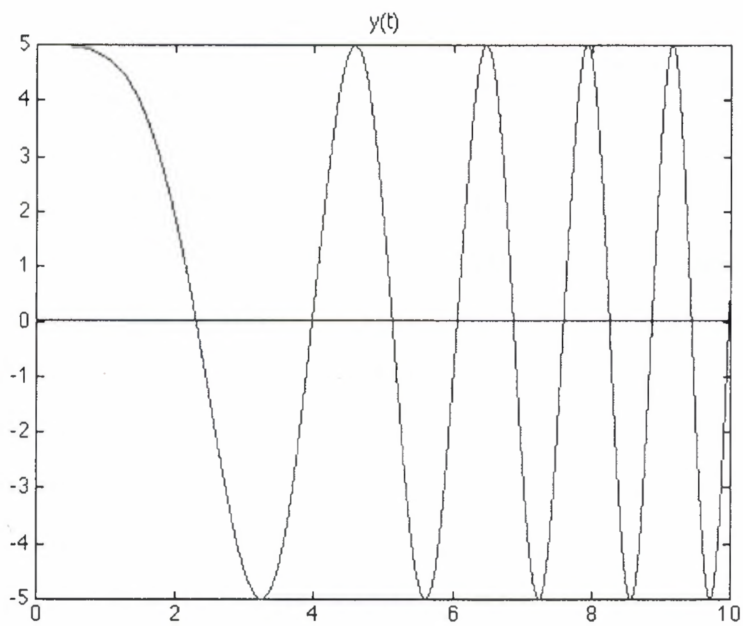


Figure: 3.20

3.2 The carrier

Amplitude modulation operates on a specific frequency known as a carrier. This signal never changes in power. The operating power of a broadcasting station is the carrier power. The radio signal is generated in the form of a sine wave. Figure below, left depicts its wavelength and amplitude characteristics. The number of wavelengths occurring in one second is the wave's frequency, measured in cycles per second, or Hertz.

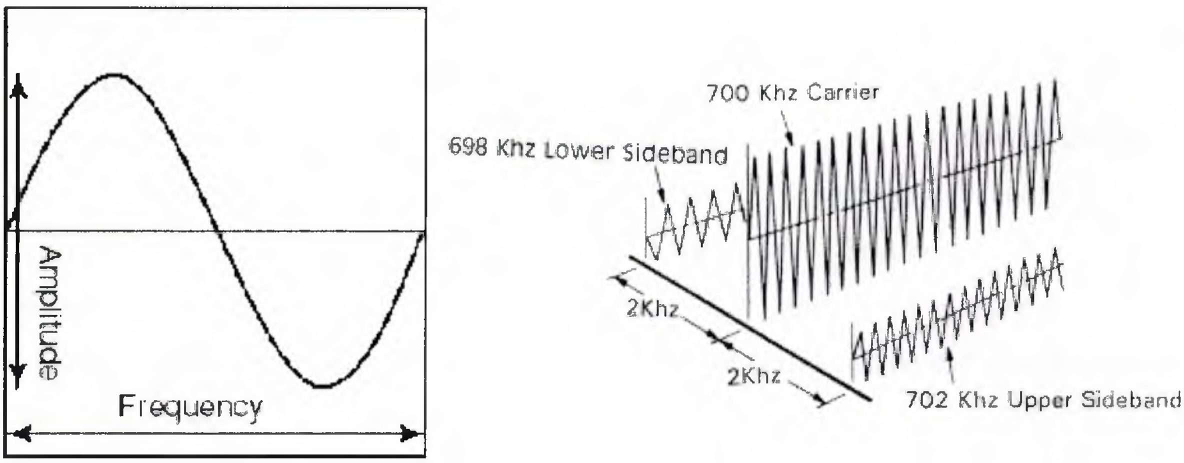


Figure: 3.21

The audio signal is added to the carrier frequency creating modulation. For instance, if the carrier frequency is 700 kHz, the radio signal, or carrier, is creating magnetic fields that are radiating off the antenna at 700,000 times per second. If audio is applied to this carrier, the sum and difference frequencies also will be transmitted. If an audio tone of 2,000 cycles is applied to the carrier, the following frequencies will be present:

Audio Frequency: 2,000 Hz

Carrier Frequency: 700,000 Hz

Sum Frequency: 702,000 Hz

Difference Frequency: 698,000 Hz

The antenna will accept the frequencies that are most closely related to the carrier: 698 kHz, 700 kHz and 702 kHz. The 698 kHz and 702 kHz are sidebands. The difference between the carrier and sideband frequencies is the audio frequency. It is duplicated above and below the carrier.

Figure 3.21 depicts the sidebands and the carrier. The amplitude of the sideband determines the loudness of the signal while varying frequencies in the sideband represent the audio information. For years, radios used a diode or envelope detector to extract the audio from the radio signal and amplify it.

Much of the problem with AM broadcast today is not within the transmission of the signal but in its reception. Most electro-magnetic noise, from lightning, motors, computers, etc., is an amplitude function. Because the AM receiver is detecting amplitude variations, it receives the desired signal along with any other electro-magnetic noise in the vicinity.

Remember the radio signal is very weak. Signals from computers, telephone systems, appliances, and so many other local sources are much stronger. The receiver picks up everything surrounding the carrier and amplifies it, often producing a lot of noise.

Over the past 30 years, receiver manufacturers have tried to reduce noise by narrowing the frequency bandwidth of the tuner.

AM transmits a frequency response that is very close to human hearing. It is flat out to 7,500 Hz and beyond. However, audio is varying constantly, allowing noise to get in where low levels of radio signal are present. The receiver manufactures decided to cut the audio bandwidth to 2,500 to 3,000 Hz. This reduces fidelity.

There are other methods available to reduce noise in AM while keeping audio fidelity high. One is to replace the envelope detector with a synchronous detector. That was once an expensive addition, but now simply requires a microprocessor. Receivers with AM stereo capability use them with good results. Many automobiles have them. They do not eliminate noise, but they reduce it.

Denon has made a receiver sold by the NAB that has a "smart filter" that will eliminate a lot of noise. However, few consumer stereo manufacturers have chosen to add this feature in their AM receivers. Even the ingenious Bose wave radio, which uses a synchronous-type of detection circuit, has no provision for decoding a stereo signal.

3.3 Strengths of AM

In the 1930s and '40s there were far fewer radio stations on the air, producing far less co-channel and adjacent channel interference. Primary service areas received a strong signal with little local interference to compete with the radio signal. Receivers were designed with radio frequency circuits that were broad enough to receive the entire signal.

Life was good for AM. I grew up in the New York metropolitan area, where there were eight 50 kW AM stations. These signals were so powerful that lightning often caused only a small crackle to the audio.

After World War II there was a demand for more broadcasting stations for smaller communities. The FCC reduced interference standards, resulting in thousands of new stations. Many were awarded licenses for daytime-only operation, while others were given full-time authority. Many had narrow directional patterns, a scenario that fit the audience of the times, far less mobile than that of today.

The addition of stations from the 1940s until today created problems for receiver manufacturers. Receivers were no longer picking up clean signals free of whistles (heterodynes) and adjacent channel audio. The solution was to narrow the bandwidth of the received signal. This has been done to the point where now AM, on some of the best tuners available, sounds like it is being received over a telephone line.

AM stereo was an ingenious idea that actually improved the sound quality of AM, giving it depth. Those who have heard it know. Unfortunately, the system never got off of the implementation stage with few receiver manufacturers making radios. I always thought a great disservice was done to AM in the way stereo development was handled by the FCC, the inventors, receiver manufacturers, and even the broadcasters. It has been an opportunity lost, in my opinion.

AM transmission equipment is still capable of doing what it did 50 years ago. Most receivers are a shadow of their ancestors when it comes to reproducing the signal. The technology is there to improve fidelity, but most consumer receiver manufacturers fail to take action. In fact, many FM receivers don't approach the quality FM broadcasters are transmitting.

Audio levels are extremely important for AM broadcasters. The strength of the power in the sidebands creates the "loudness" of the signal in the receiver. It is important to have a strong signal to overcome as much noise as possible.

Audio levels are observed in the studio by monitoring the VU meter. Located on the audio console, this volume unit meter measures the electrical strength of audio signals being broadcast.

The console receives signals from microwave remote broadcasts or satellite services as well as from the studio and combines or chooses them from local broadcasting. The VU meter measures signal strength in decibels. This a logarithmic measurement that responds to sound the same way our ears do. The '0' level on the VU meter is the optimum operating level.

Audio quality starts in the studio. To reproduce audio properly, it is important to have the VU meter readings peak around 0 or +1. Because audio varies constantly, the signal level should be monitored to make sure that the loudest audio is at this level. Do no listen with your ears; observe with the meter. Monitor speakers can be deceiving; the audio system can be checked with calibrated signals from audio oscillators, test CDs and tapes providing a '0' level to make sure the console is operating properly. Along with attenuators that adjust the signal, consoles also have trim adjustments to make sure levels are accurate. Trim adjustments also allow for a balanced output of stereo consoles.

It is important to make sure that the material recorded for broadcast is prepared in a proper manner. If audio is recorded improperly in either a digital or analog format, problems will occur. Recording at an insufficient level will permit the introduction of noise. Recording at an excessive level of about +1 dB will cause distortion and loss of dynamic range, the ability to capture audio in its proper range. Every source of audio that emanates from the console should produce the same peak levels, giving the listener a "feel of continuity" from one segment to the next.

If the station is broadcasting in AM stereo, make sure the channel phasing is correct. Most stereo consoles include left, right and sum or mono VU meters. If you receive material recorded out of phase, or a problem occurs within the studio that causes the left and right channel to be out of phase with each other, the signals that are common to both channels will be canceled out. That will cause problems for your listeners, because most AM receivers are monophonic.

In most instances, a phase problem will cause significant loss of signal because the majority of any stereo signal has components common to both channels. When signals are out of phase, the left and right meters of a stereo pair will appear to read normally, but the mono meter on your mixer will drop to zero.

VU meters can also be used to trace hum and noise. For example, if a VU meter will not return to its resting place, it may indicate an unwanted signal in the system. This can be traced by removing audio from the console and turning up the audio monitor. Increasing and decreasing the levels of individual channels along with cutting off inputs can pinpoint the location of the problem; it can be within the console or an external source.

Another important instrument, though not required, is the modulation monitor. It allows you to see what the signal looks like after it has been transmitted, measured in percent as well as dB. It should be as readily accessible as the console's VU meters.

CHAPTER IV: APPLICATION OF AM

4.1 Audio Preparations for AM Radio

Audio is a variable: It constantly changes and creates the power levels in sidebands. Because noise is an amplitude function, it is important to create a power level that gives the best signal-to-noise ratio that can possibly be attained.

Audio can come from any number of digital or analog sources. It must be delivered to the transmitter at a level that will give the maximum sideband power possible without adding distortion or reducing dynamic range to an intolerable level. Overdriving the signal at the studio console can cause clipping distortion. Once distortion is introduced into the system, it cannot be removed.

After the signal leaves the studio, it travels either across the building to the transmitter or through a studio-transmitter link (STL), telephone line or microwave feed to the transmitter location.

The audio is prepared for the transmitter by using compressors or limiters to prevent distortion and overmodulation. Often it is wise to use a limiter on telephone lines or microwave STLs to prevent distortion. Unlimited instantaneous peaks can cause undesirable affects.

All transmission systems have a specific operating level, usually related in decibels, which should not be exceeded. The amount of protection necessary is dictated by the type of format being used. Conversations with the engineers in charge of the station can assist you on arriving at a level that is appropriate for your station. If your station does not sound as good as others in your market, assume something is wrong.

4.2 Masking noise

Audio processing developed rapidly in the 1950s. After television took away many radio audiences, broadcasters adopted the DJ format, becoming music or news stations. To hold onto listeners, they made the station appear loud. Limiters and compressors were introduced to increase loudness at the expense of some dynamic range, which reduced the appearance of noise within the system.

Another technique to increase loudness and coverage area was the introduction of reverberation. Many stations put these devices at their transmitter sites, adding an echo effect to the broadcast signal. This gave the DJ a louder and more commanding voice. One might say everything sounded like it was happening in a large bathroom, but it worked. Music performers - such as Phil Spector - began to add echo to their songs. Spector's "wall of sound" recordings were made with the AM band in mind.

The sound of an AM broadcasting station is dependent upon everything from the quality of the studio console to the condition of the antenna system. AM is frequency-sensitive; sideband power is directly related to the audio frequency creating the modulation.

The most powerful sidebands are closest to the carrier. Check the frequency response of the audio line, telephone line or STL. Don't assume everything is okay. There was a time when the FCC required this in the form of the audio proof-of-performance. A flat frequency response in the audio line can translate into a louder audio signal in the receiver.

Format plays an important part when considering the amount of audio processing being used. You also have to consider what type of receiver people will be listening on. AM listening takes place in cars probably more than any other place. It is important to prepare a signal that will be able to compete with all of the other sounds around the vehicle.

4.3 Chemistry set

Processing must not be to the point where "listener fatigue" is reached. Compressors and limiters can be set so the modulation meter almost constantly remains at 100 percent. While this will produce an audio signal almost void of noise, it will also annoy the listener. It usually takes some experimenting with an audio system to get it precisely where you want.

Classical music, which contains many audio levels and variations, requires little audio processing. For that reason, classical music is difficult to carry on AM in this age. At times, the classical station listener might think the station is off the air. Popular music is different and can be processed at a more aggressive level, as its dynamic range is much smaller.

For talk radio, audio should be set somewhere between what is best only for voice and music. Commercial material may include jingles and short music pieces, and they must sound right when broadcast. The "attack time" of the compressor/limiter should not be set to increase the audio output level at the instant someone stops talking. It is important to read the manuals on proper setup of these units to achieve the sound you and your listeners want.

The compressor/limiter has become quite a sophisticated piece of equipment over the years with features that include audio compression, expansion, limiting, clipping and gating. Many models are multiband, permitting specific processing at prescribed frequencies.

When choosing an audio processing system, know that all systems do not interface favorably with each other. A certain manufacturer may say its processor will improve your audio with documented evidence from other broadcasters. While this may be true, you still need to test a unit in your audio chain. Install a demonstrator unit according to the manufacturer's instructions and listen to it under differing conditions: in a car, on a Walkman, at home, etc. Make sure it is performing to your expectations.

An equalizer may be used to enhance certain frequencies, but this device really "unequalizes" your signal by lowering the output of some frequencies while increasing others. Some broadcasters attempt to use an equalizer to improve the frequency response of an AM station that has a deteriorating antenna system. This is not recommended. The equalizer is best used in the production studio to improve the sound of poorly recorded audio, perhaps to remove unwanted hum from a remote feed. It is not recommended as a part of the studio to transmitter audio chain.

I am probably the last person in America to say this, but I will: AM stereo is a way to improve the sound of AM. I recently researched AM formats and found that more than 75 percent of national AM stations program a substantial amount of music. AM is not the news/talk world that seems to dominate the major markets. AM stereo is worth the investment

4.4 Using a Modulation Reactor

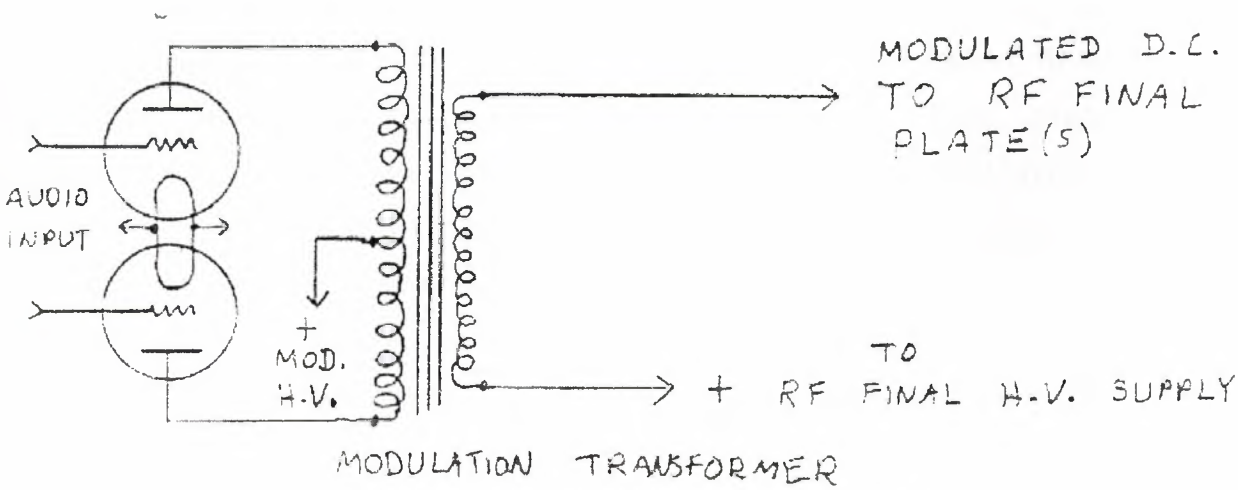
One big difference in AM operation on the bands today compared to earlier days is the widespread interest in high fidelity audio. Back when AM was the dominant mode on the amateur phone bands, relatively little attention was paid to audio fidelity. The important thing was "communications quality", usually with restricted frequency response, in hopes of achieving better "penetration". In addition, audio quality was often further degraded by the use of primitive processing techniques such as hard clipping followed by simple lowpass filters with inherently severe phase shift characteristics.

Nowadays, AM users tend to take pride in what their signals sound like. There is little likelihood that amateurs obsessed with "penetration at any cost" would operate AM; their ideal mode came on the scene with the advent of SSB. AM users today use transmitters and speech equipment with quality equal to or exceeding that of professional broadcast equipment. Processing, if used, is frequently accomplished with sophisticated studio quality devices retired from the broadcast service, if not homebrewed. The ideal AM signal has broadcast station sound, yet is able to penetrate the QRM and QRN normally heard on the amateur bands, and cope with adverse conditions such as selective fading. This is never fully accomplished, but striving for this ideal has become one of the popular facets of amateur AM.

AM transmitters heard on the amateur bands include military surplus, commercial communications transmitters, commercially built amateur transmitters, homebrew amateur transmitters, and broadcast transmitters. Except for the latter, most of these transmitters are designed for communications quality audio. Some homebrew ham rigs have been built to broadcast standards, and many other AM users strive to accomplish this result.

With plate modulated transmitters, achievable audio quality is largely determined by the audio transformers. Commercially built and older homebrew ham transmitters are usually equipped with very low quality "amateur grade" transformers, reflecting "economy" and a longstanding attitude in amateur circles that audio quality is totally unimportant and that the only legitimate concern in amateur voice transmission is whether or not the signal is copyable at the other end of the QSO. Commercial and military rigs are usually a step above "amateur grade" audio, but signal readability is still the overriding concern in their design. Nevertheless, many of these transmitters, built with "commercial grade" audio transformers, sound quite good when a high quality microphone is used.

Apart from the quality of the audio transformers used, the circuit arrangement for voice communications transmitters and high fidelity transmitters is similar, although the coupling circuits in the communications transmitter may be deliberately designed to restrict the frequency response to something on the order of 300-3000 hz. However, there is one important difference in the way the modulation transformer is connected in broadcast transmitters compared to communications transmitters. In communications transmitters, including amateur rigs, the secondary winding of the modulation transformer usually carries the full DC plate current to the final amplifier stage. This familiar circuit is shown in the Figure 4.1.



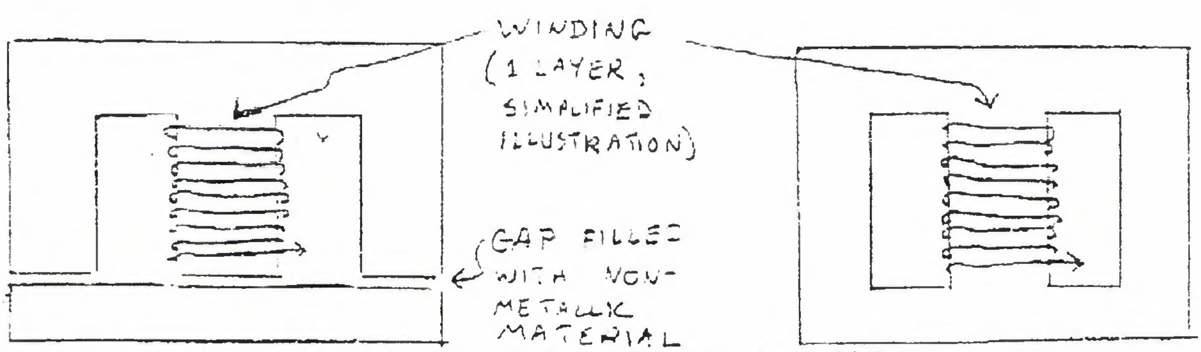
Conventional modulation transformer hookup

Figure 4.1

This circuit has the disadvantage that the final amplifier direct current flowing through the secondary winding of the transformer generates a "magnetic bias" on `the transformer core, and reduces the effective inductance of the transformer windings. This works exactly like a swinging choke in the power supply, wherein the DC flowing through the winding reduces the amount of inductance. To avoid magnetically saturating the core of the transformer, the laminations of the modulation transformer are stacked in such a way that there is a gap in the core, in exactly the same manner that power supply chokes are constructed.

This gap reduces the tendency of the core to saturate, but it also reduces the effective amount of core material in the transformer and thus reduces the inductance of the winnings. Furthermore, the gap in the core does not totally eliminate the DC saturation. Therefore the plate current flowing through the winding reduces the inductance even further. The low frequency response of an audio transformer is directly related to the inductance of its windings, so the core gap and direct current flow reduce the low frequency response of the transformer. For "communications quality" audio this is not considered serious and therefore most non-broadcast transmitters are designed to allow the plate current to flow directly through the modulation transformer secondary. For high fidelity audio, the low frequency attenuation caused by this arrangement is so severe that a tremendously oversize modulation transformer would be required. Other problems inherent to this scheme include phase shift distortion due to the restricted low frequency response of the transformer, and distortion caused by near-saturation of the transformer core on audio peaks, not to mention heating of the transformer winding due to the current flow. The end result is that the transformer often leaves a muddy sound on the audio and may run quite warm. The phase shift can greatly reduce the effectiveness of certain audio processing techniques, particularly the simple speech clipping often used in amateur AM transmitters. An additional problem sometimes associated with modulation transformers which carry plate current is a tendency to "talk back". This is annoying to the operator and can degrade audio quality by generating acoustical feedback through the microphone. The popular UTC VM-5 and CVM-5 transformers are notorious for this.

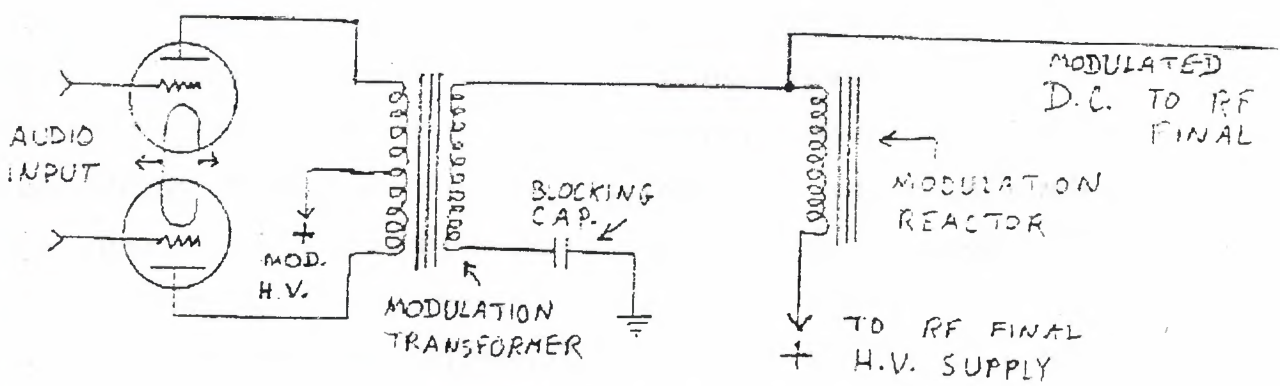
These problems are eliminated in broadcast transmitters by isolating the DC plate current from the modulation transformer winding, allowing the modulation transformer to only carry audio. Since there is no DC flowing through the winding, the transformer core can be stacked just like an AC power transformer, without a gap. Such a transformer makes much more efficient use of the iron in the core, giving as much as ten times the inductance per winding compared to a similar size transformer with a gap in the core. The result is much better low frequency response for a given core size. In addition the transformer runs cooler, generates less overall audio distortion, is less prone to acoustical vibration, and the likelihood of accidental transformer burnout may even be reduced.



Core with gap (left), without gap (right)

Figure: 4.2

The DC is isolated from the modulation transformer secondary in broadcast transmitters by adding two additional components to the circuit. A blocking capacitor is placed in series with the modulation transformer secondary. A value of capacitance is selected to cause negligible rolloff at the lowest audio frequency within the specified response of the transmitter. To carry the plate current, a choke is connected between the final amplifier and the DC plate supply. This inductor is commonly called the modulation reactor. It is effectively placed in parallel with the modulation transformer secondary, and therefore must have sufficient inductance to not affect the low frequency response of the transmitter.



Modulation transformer hookup using blocking capacitor and modulation reactor

Figure: 4.3

In a typical plate modulated transmitter, the DC blocking capacitor will have a value between 1 and 10 microfarad. The lower the modulating impedance (final amplifier plate voltage in volts divided by plate current in amperes), the higher the capacitance required for a given low frequency rolloff. Many of the older broadcast transmitters had excellent low frequency response with blocking capacitors as small as 2 mfd. Of course, more capacitance than needed, within reasonable limits, won't hurt anything. High voltage oil capacitors in this range are still available at flea markets at very low cost, since this capacitance is insufficient for most modern power supply filters. A good value to use is 4 to 8 mfd., which will work fine with about any modulating impedance encountered in a normal transmitter at amateur power levels. Since the high voltage to the RF final appears across the capacitor, the minimum DC working voltage should be at least 1.5 times the highest unmodulated plate voltage expected to be applied to the final amplifier, to be on the safe side.

The modulation reactor is the most difficult to find item required for this circuit. A typical reactor is rated at somewhere between 25 and 60 Henries at the maximum final amplifier plate current under normal operation. This is several times the inductance of a typical power supply filter choke. Amateurs have successfully used 10 Henry smoothing chokes as modulation reactors, although such a low inductance will reduce the low frequency response of the transmitter. One solution is to wire up several power supply filter chokes in series, although this tends to take up a lot of space and is somewhat of a “JS” setup. The preferred component to use is a real modulation reactor designed to go in a broadcast transmitter. With so many tube type AM broadcast transmitters going out of service at this time, these can be found if you are willing to look for them.

4.5 Circuit Variations

Figure 4.3 shows the most common modulation reactor circuit used in broadcast transmitters. In this circuit, the blocking capacitor is placed between the "cold" side of the modulation transformer secondary and ground. This provides a direct round return for the audio, independent of the high voltage power supply. In the conventional circuit (Figure 4.1] , the audio is returned to ground through the hv power supply. The output filter capacitor is effectively in series with the modulation transformer secondary, and unless sufficient capacitance is used, this will limit the low frequency response of the modulator. More seriously, any residual hum in the output of the power supply will modulate the final amplifier and be audible on the signal. With the circuit in Figure 4.3 the hum output from the power supply is not placed in series with the modulator, and furthermore, the modulation reactor serves as an additional high inductance smoothing choke. Thus, the circuit in Figure 4.3 results in a substantial reduction of the hum level of the transmitter.

There is yet another advantage to this circuit if a common power supply is used for the class B modulator and final. Because of instantaneous variations in the load presented to the power supply by the modulator, a strong harmonic distortion product appears at the centre tap of the primary winding of the modulation transformer. With the conventional (Figure 4.1) circuit, the "cold" side of the modulation transformer secondary is tied directly to the centre tap of the primary winding. Any harmonic distortion products existing at that point appear in series with the modulator and thus add distortion to the audio, especially at low audio frequencies. This problem can be reduced by using a large filter capacitor at the output of the power supply, or by using an additional section of filtering, consisting of a smoothing choke and another filter capacitor, between the centre tap of the modulation transformer primary and the “cold” end of the secondary. However, with the circuit in Figure 4.3, such precautions are unnecessary because the modulation transformer secondary is completely isolated from the RF final plate supply as far as audio is concerned.

One disadvantage to the Figure 4.3 circuit is the possibility of modulation transformer failure due to a high voltage arc between the primary and secondary windings of the transformer. The "hot" end of the modulation transformer secondary is tied directly to the high voltage lead to the final, so that the secondary winding remains at the DC potential applied to the final amplifier plate. But the "cold" end is connected to ground through the blocking capacitor. When the high voltage is turned off, as for example when the transmitter is in standby, the blocking capacitor discharges along with the power supply filters; when high voltage is reapplied, the modulation transformer secondary remains momentarily at ground potential until the capacitor becomes recharged, through the combined inductances of the modulation reactor and modulation transformer secondary. During this brief instant the full DC potential of the modulator plate supply appears between the two transformer windings. This transient may be sufficient to cause an arc to occur and destroy the insulation between the windings. This is unlikely to occur with a transformer that has been maintained in a dry environment, since most modulation transformers are designed to withstand a substantial voltage difference between windings. Most broadcast transmitters use the circuit shown in Figure 4.3, and communications transmitters may use separate power supplies for the modulator and final with a substantial difference in output voltages. It is possible, however, that some modulation transformers may be designed specifically for a transmitter wherein the windings always remain at the same DC potential, and thus the insulation between windings may be prone to breakdown. Fortunately, this does not seem to be a problem encountered very often.

Figure 4.4 shows some variations of the basic circuit shown in Figure 4.3. In Fig. 4A, the "cold" end of the modulation transformer secondary is tied to the high voltage supply as in Figure 4.1, and the blocking capacitor is inserted in series with the "hot" end. The circuit in Fig. 4B is identical to 4A, except for the placement of the capacitor, which has been switched to the "cold" end. Both these circuits function much in the manner of the conventional (Figure 4.1) circuit, but the DC is blocked from flowing through the secondary winding. With this arrangement, several of the advantages mentioned with Figure 4.3 are lost, because the audio is returned to the high voltage plate supply as in the conventional circuit. However, the danger of modulation transformer breakdown described in the previous paragraph is eliminated.

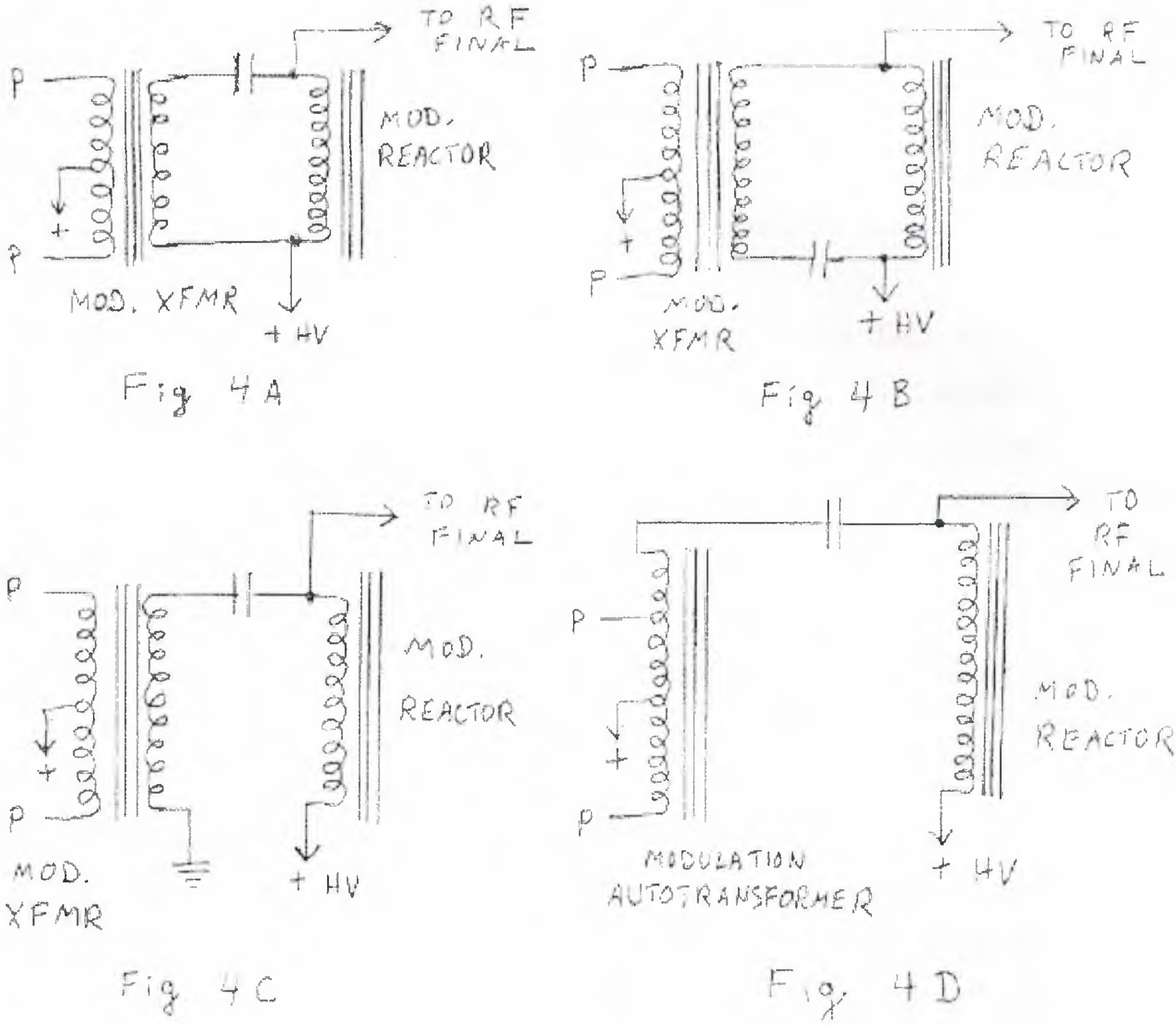


Figure: 4.4

The circuit in Fig. 4.4 (4C) is not recommended, since the modulation transformer secondary is tied directly to ground, and the full DC plate voltage to the modulator appears between the windings at all times, and as explained previously, could cause breakdown of the modulation transformer. Nevertheless, some commercially built AM broadcast transmitters use this circuit even though it presents no advantage over the Fig. 4.3 circuit. Fig. 4D shows how a reactor may be used with a modulation autotransformer in which there are no separate primary and secondary windings. This circuit offers the previously mentioned advantages of blocking DC from flowing through the modulation transformer winding, and allows the option of using separate power supplies for the modulator and final amplifier with a modulation autotransformer. However, the audio output from the modulation transformer is returned through the modulator high voltage plate supply, necessitating the use of a large value of output capacitance in the filter.

In Figures 4.4 4A, 4B and 4D, there is no DC voltage across the blocking capacitor. Theoretically, DC working voltage should be unimportant. However, transients and very low audio frequency voltages may appear across this capacitor under a variety of conditions, so the safe bet is to always treat this capacitor as if the full DC to the final appeared across it, and use the same minimum working voltage as required in Fig. 4.3.

4.6 Which Modulation Transformer to Use with a Reactor?

As previously explained, modulation transformers designed for use with a reactor have no gap in the core, and are not designed to allow DC plate current to pass through the secondary. However, transformers designed to carry DC through the secondary can be used with a reactor, and improved performance can be expected. In spite of the core gap, the low frequency response will be improved when the transformer is not called on to carry DC. The transformer will run cooler without the DC, and audible vibration will be reduced. I remember using and old UTC VM-5 which talked back terribly without the reactor, but it was practically silent when the reactor was wired into the circuit. Mike, NI4N was pleasantly surprised at the improvement in audio quality when he installed a reactor in the 1938 vintage transmitter built by his grandfather, which uses an old bottom-of-the-line Thordarson "ham radio" quality modulation transformer. In addition, it is likely that the power rating of a non-broadcast modulation transformer can be safely increased somewhat beyond the manufacturer's recommended maximum when a reactor is used. If a good broadcast quality modulation transformer designed to be used with a reactor is available, by all means use it. Otherwise, expect better performance with a conventional "ham radio quality" transformer when the reactor is included in the circuit.

If you are ambitious, you may be able to convert a gapped core to a solid core by restacking the laminations, sometimes referred to as "crosslaminating" a transformer. This is quite easy to do if the transformer is the open-frame type, and no potting compound is used on the winding. If the transformer is impregnated with varnish, it will have to be carefully heated to a temperature that will soften the varnish, but not damage the transformer insulation. If you are really ambitious, potted transformers can be heated, the tar removed, and then repotted after the modifications are made. Cross-laminating will greatly improve the low frequency response. However, it is recommended that you have some solid hands-on experience at assembly and disassembly of transformers, and all the necessary tools and workshop facilities, before attempting to tear apart something as hard to find these days as a good modulation transformer.

Thus ends Part I of this article. Next issue we will conclude with data on how to determine the amount of inductance required for optimum performance with a modulation reactor, and we will include catalogue data giving the manufacturers' ratings of modulation reactors produced by several prominent transformer companies.

4.7 Amplitude Modulation for Television and Monitors

The cable television systems are a clear example of multiplexing where electricity rather than light is used to carry a multitude of signals to your home simultaneously. Cable television also differs from our Alice and Bob example in that the signal being multiplexed is not digital. Until very recently, all the television signals entering homes were encoded in analog form. Surprisingly, despite the emphasis we have placed on the difference between analog and digital encodings, the techniques used to support multiplexing of digital and analog signals are not that different. Rather, the techniques used in the digital world can be seen as a special case of a general technique called modulation that can be applied in both domains. In this section we will discuss this general technique.

We will use sound as an example of a form of analog information that can be encoded using modulation. While cable systems use modulation to encode a complete television signal, such a signal is more complicated than necessary for our purposes. Sound, whiich is of course a component of a television signal, is sufficient to illustrate the principles underlying modulation and it is considerably simpler to discuss than television signals.

One of the oldest applications of modulation is to encode sound waves in radio transmimssions. While you may think you have never heard of modulation before, you probably use abreviations that refer to modulation techniques quite regularly. Your stereo, your car radio and your alarm clock radio probably all have switches that lets you choose between AM and FM. AM stands for "Amplitude Modulation" and FM stands for "Frequency Modulation." These are the names of two commonly used modulation techniques. There are others. Fortunately, our goals will be satisfied but just discussing one of these approaches, amplitude modulation.

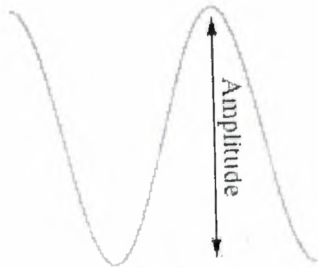


Figure: 4.5

The amplitude of a wave is size of the total variation in whatever physical quantity varies as the wave passes by. For a water wave, the amplitude is the distance from the highest point of a peak to the lowest point of a trough. For an electromagnetic wave, the amplitude is the size of the variations seen in the electrical or magnetic forces as the wave passes by. As the diagram on the right indicates, when we plot a wave, the amplitude determines the height of the plot. In the case of light, the amplitude will be proportional to the brightness of the light. The brighter the light, the larger the amplitude. In my dictionary, modulate is defined as "To adjust or adapt to a certain proportion." The basis of amplitude modulation is to adjust the amplitude of waves as a way to encode information for transmission.

The information they had to encode was a digital signal composed of sequences of 1's and 0's. They transmitted these signals by turning on their signal lights whenever they wanted to send a 1 and turning their signal lights off whenever they wanted to send a 0. While this flashlight example captures much about the techniques actually used to multiplex signals between devices, it ignores one important detail, speed. No matter how good their coordination might be, Alice and Bob could not turn their flashlights on and off rapidly enough to match the transmission speeds of current device systems. A very common transmission rate used on optical fiber is 100 million binary digits per second!

In an actual communications system, the flashlights are replaced by light emitting diodes or lasers and the manual switches on the flashlights are replaced by switches that are electricly controlled. An electrical signal encoding the binary sequence to be transmitted is generated by a device. This signal is then used to control an electronic switch that turns the signal light on and off. The figure 4.6 below illustrates this arrangement.

Figure: 4.6

When there is current arriving from the device, the electronic swtich is turned on and sends power to the light source. As the light flashes on and off, the pattern of light waves produced encodes the same sequence of 0's and 1's encoded in the electrical signal that the device sends to the electronic switch. The signals leaving the bulb are light waves, so our diagram shows pulses of waves separated by quiet periods emanating from the light bulb. The only inaccurate aspect of the waves shown is that the frequencies of light waves are typically over a million times greater than the rate at which devices send binary sequences. So, while each pulse in our diagram is wide enough to hold just one or two osciallations of the wave pattern, within a pulse of light in a real transmission system there would be millions of wavelengths of light.

Now, consider how to send a signal representing a sound wave. While certains sounds, like a single flute playing a single note, produce nice periodic waves, most real sound are composed of waves in which the shape and amplitude of each peak is slightly different from all the other peaks. A snippet of what such a sound wave might look like is shown in figure 4.7

Figure: 4.7

This waveform serves as blunt reminders of the difference between analog and digital forms of information. Compare the smooth transitions seen in this waveform to the rectangular pulses seen in the digital signal sent from the device to the electronic switch in the preceding diagram. While the transmission of digital information only requires a small, finite set of signals -- in this case two -- analog signals vary continuously. For any two points along the sound wave there are infinitely many "in between" points. We need a way to transmit information about all these minute variations.

To transmit analog signals, a simple on-off switch is not sufficient -- even if it is an electronic one! To address the range of variations possible in an analog signal we need something more like a dimmer switch. Luckily, we can construct an electronic dimmer switch, that is a dimmer switch which takes one electical signal as its input and determines how much power to output based on the strength of its input at any moment. A microphone can then be used to convert a sound into a varying electrical signal that can be used to control the electronic dimmer switch. Finally, the electricity flowing out of the switch can be attached to a light source.

The figure 4.8 shows how this system to encode sound for transmission through optical fiber would work.

Figure: 4.8

One can see that the amplitudes of the light waves used to transmit the signal are being varied in such a way that their amplitudes are always proportional to the value of the input signal at a corresponding point in time. We make this more obvious in the figure 4.9 below by overlaying the graphs of the electrical signal from the microphone with that of the forces associated with the light waves.

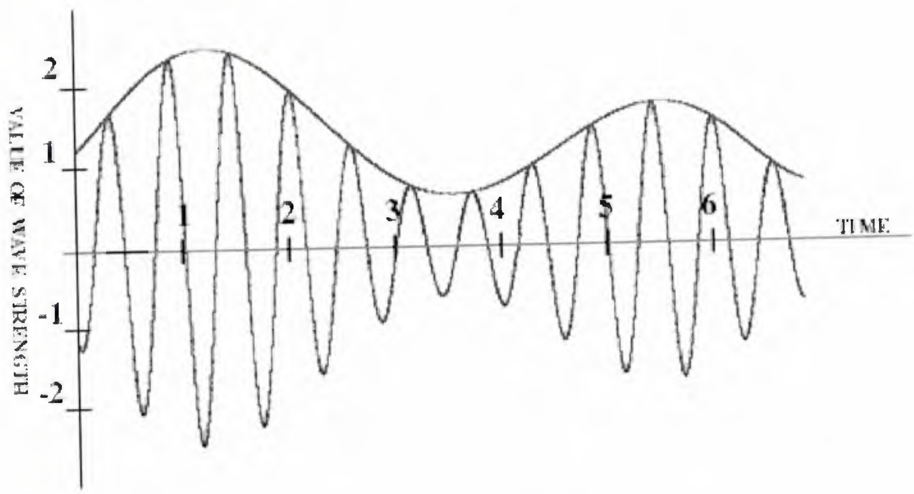


Figure: 4.9

Looking at this picture, we can see that the graph of the light wave is clearly related to the graph of the audio signal, but the relationship is not simply equivalence or proportionality. In fact, the value of the light wave at any particular point has little to do with the value of the audio signal. Only where the light waves peak do the graphs of the two waves meet.

To fully appreciate what is going on here it is important to distinguish between the value associated with a wave at a point and the amplitude of the wave. The value is a single measurement of the quantity varyied by the wave at a particular point and time. For example, at the points labeled 5 and 6 on the horizontal axis of the graph, [3] the values associated with the red curve representing the audio signal are both just about 1.5 on the vertical scale. The values of the purple curve representing the light waves differ significantly at the points 5 and 6. At 5, the value of the purple wave is just about 0 while at 6 it is 1.5. On the other hand, if one were to ask about the amplitude of the purple wave near the points 5 and 6, the answer in both cases would be "about 3". When estimating the amplitude, one doesn't just look at the value at the point specified. Instead, one looks at the behavior of the wave for about the duration of one period centered on the point. Near both points 5 and 6, the light wave's value varies from a high point of 1.5 to a low point of about -1.5. Accordingly, in both regions, the amplitude or total variation of the wave is about 3.

The relationship between the two waves can be expressed fairly simply in terms of amplitude. The effect of the electronic dimmer switch is to ensure that the amplitude of the light wave is always proportional to the value of the audio signal.

Since "proportional" is certainly a mathematical term, it would be nice if we could go one step further and express the relationship between the two waves shown in our diagram mathematically. First, we need some mathematical notation to let us talk about the two waves themselves. From the plot, we can see that both waves are functions of the distance from the source. To start, let's give these functions names. We will call the audio signal "a" and use the standard notation a(t) to denote the value of the audio signal at some particular time t. Similarly, we will use the letter "s" as a name for the signal actually transmitted through the fiber and use s(t) to denote the function that corresponds to the strength of the forces in that signal at any time t.

The next step in seeing how to express the relationship between a(t) and s(t) is to imagine what s(t) would look like for an unrealistically simple example of an audio signal. Sound waves are all about variation in air pressure. Suppose, however, that for some period of time in a particular sound wave, the pressure stayed constant. The graph of such an audio signal together with the graph of the signal that would be transmitted if such a constant value audio signal was fed to the electronic dimmer switch to produce a light beam for transmission would look like in figure 4.10:

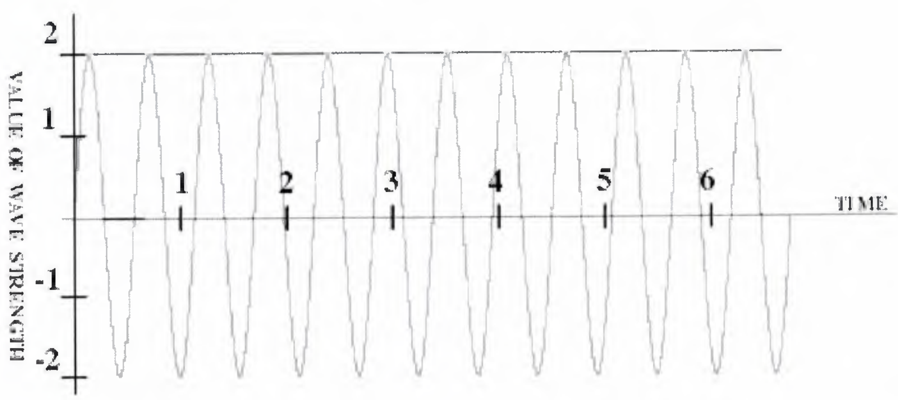


Figure: 4.10

This would be boring if it lasted for very long, but it is simple enough to help us get some insight into the relationship between more complex audio signals and the transmitted signals that would be used to represent them. The transmitted signal in this graph is clearly periodic. Each repetition is identical to all the others in wavelength and in amplitude. There are several simple mathematical functions that are periodic. One of the simplest is the trigonometric function "sine". Assuming we measure angles in radians instead of degrees (1 radian = $180/\pi$ degrees), the graph of the sine function, sin(t), looks like figure 4.11:

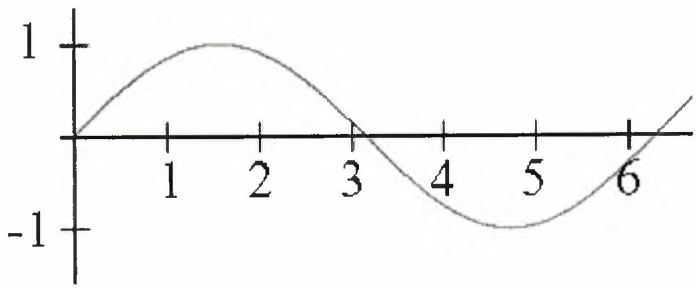


Figure: 4.11

While both the transmitted signal shown in the preceding figure and the sine function are periodic, they are clearly not identical. The transmitted signal looks like a version of the sine function that has been squeezed horizontally and stretched vertically. We can perform this stretching and squeezing mathematically to obtain a mathematical formula that describes the transmitted signal, s(t), associated with the audio signal a(t) = 2.

First, lets think about how to squeeze the sine function horizontally. As shown in the diagram, one period of the sine function stretches from the origin to the point where t = 2 x π.. If instead of graphing y = sin(t), we graph y = sin(2t), we will reduce the length of a single period by a factor of two, doubling the frequency. The first period of this new function stretches from the origin to the point where t equals π..

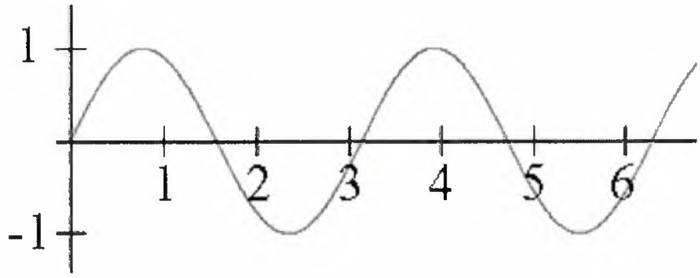


Figure 4.12

We can squeeze the function even more by multiplying t by a number larger than 2. An interesting choice is to multiply by 2π.. The resulting function, $y = \sin(2\pi t)$ has both a wavelength and a frequency of exactly 1.

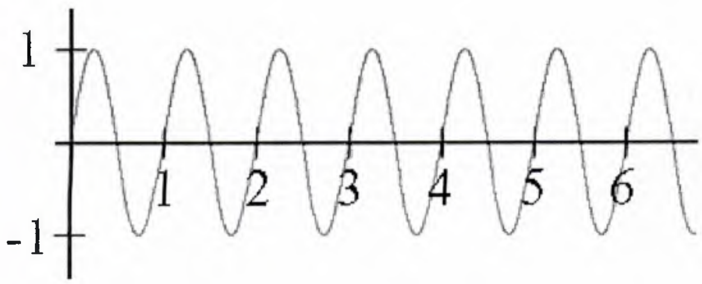


Figure 4.12

We have already seen that if we replace sin(t) by sin(2t), the frequency doubles. Similarly, if we replace $\sin(2\pi t)$ by $\sin(2 \times 2\pi t)$, the frequency will double. Since the frequency of the function $\sin(2\pi t)$ is 1, the frequency of $\sin(2 \times 2\pi t)$ will be two. In fact, for any number f, we can see that the frequency of the function $\sin(f \times 2\pi t)$ will be f.

Given this general mechanism for describing a sine wave of a given frequency, we can look back at the graph showing a(t) = 2 and the associated transmitted signal, s(t). The frequency of the transmitted signal shown in the diagram falls somewhere between 1 and two. The first cycle clearly ends before 1 while the second ends slightly after two. In fact, the frequency turns out to be just about 1.77. While

1.77 is a lovely number, we have already mentioned that the units used to mark the scale in our graphs is quite arbitrary. If we chose time units of a different size, 1.77 would be replaced by some other number. So, rather than discussing sin(1.77 x 2πt), we will work with the more general form sin(f x 2πt) assuming that `f' is the frequency of the light emitted by the source that produces the signal for transmission.

So, if we plot a(t)=2, s(t) and sin(f x 2 πt) where f is the frequency of s(t), the resulting graph will look like figure 4.13:

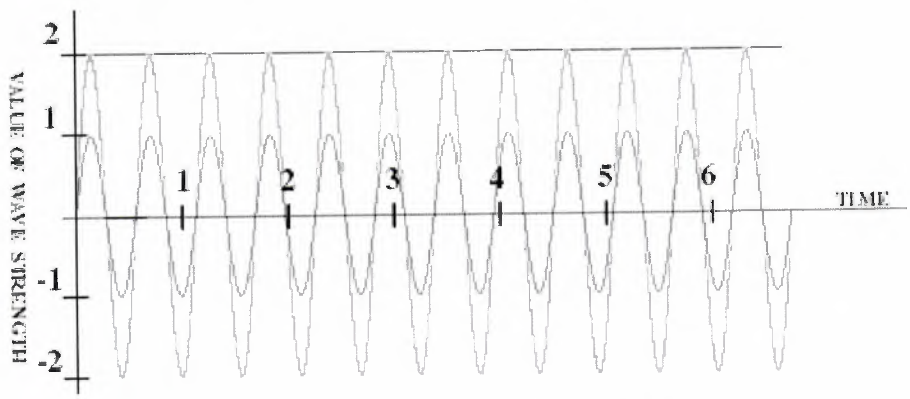


Figure: 4.13

All we have to do to make sin(f x 2 πt) match up exactly with the transmitted signal is to stretch it vertically. This is fairly easy. If sin(f x 2πt) ranges from 1 to -1, then for any number R, R x sin(f x 2πt) will range from R to -R. In the case of a(t) = 2, we want our signal to range from 2 to -1. So, a mathematical formula for s(t) is given by:

s(t) = 2 sin(f x 2πt)

In this particular case, it is clear that we could instead write

s(t) = a(t) x sin(f x 2πt)

since a(t) = 2. Fortunately, even when a(t) is not a constant, this general formula describes the relationship between the transmitted signal, s(t), and the audio signal we wish to transmit, a(t). For example, if a(t) is the function described by the red curve in the figure below then the purple curve describing the light waves in the transmitted signal is described by $s(t) = a(t) x \sin(f \times 2\pi t)$.

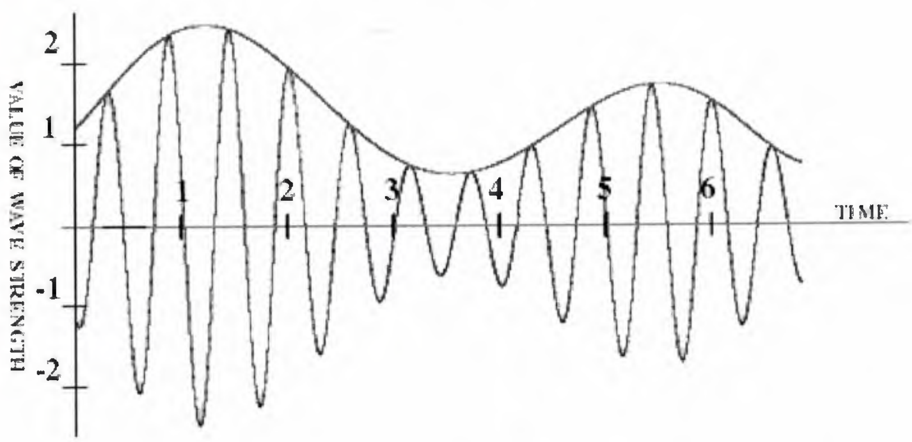


Figure: 4.14

Basically, when we multiply $\sin(f \times 2\pi t)$ by the value of a function that varies with time, we are describing a wave that oscillates with frequency f, but is not quite periodic because its amplitude varies continuously. If we multiply $\sin(f \times 2\pi t)$ by a(t), we are describing a wave whose amplitude will always be proportional to the value of a(t). This is exactly how we want the transmitted signal to behave.

Recall now that we are trying to understand a general technique for encoding information called amplitude modulation. As stated earlier, a dictionary definition for the word "modulate" is "To adjust or adapt to a certain proportion." Moduate, in this sense, is a transitive verb. That means if we are modulating, then there must be something that is being modulated. From the mathematical formula we have concocted to describe s(t), we can now identify the object being modulated. It is the simple sine wave sin(f x 2π

t). In the actual transmission system described, this wave does not actually exist. It is the wave that would be produced if we connected the light source in our imaginary transmission system to a constant power supply instead of to a varying "dimmer switch". In any case, it is important enough to the process of transmission based on modulation that it is given a name. The wave being modulated is refered to as the carrier wave. If we use the symbols c(t) to denote this wave we can then say that

$$s(t) = a(t) \times c(t)$$

where $c(t) = \sin(f \times 2\pi t)$. This bit of mathematics lets us abstract the fact that the carrier wave is a sine wave and concentrate on the facts that are most basic to understanding modulation. The basis for modulation is that a transmitted signal is produces by modifying some carrier wave in such a way that the difference between the amplitude of the unmodified carrier and the amplitude of the actual transmitted signal encode a third signal, in this case a(t).

The best part about this mathematical description of modulation is that it describes all uses of amplitude modulation, not just our imaginary fiber voice transmission system. For example, an AM radio station does not transmit using light waves. Instead, it transmits radio waves from a large antenae. For example, WABC has been broadcasting signals to the New York City area from an antenae located in Lodi, New Jersey since 1953. If you are in the New York area, you can listen to WABC by tuning your radio to 770 on the AM dial. This means that WABC transmits using a carrier wave frequency of 770 thousand cycles per second. If a(t) is the audio signal to be broadcast, the actual radio waves generated by the WABC antenae are described by the formula $a(t) \times \sin(770000 \times 2\pi t)$.

This formula is even sufficient to describe both digital and analog applications of amplitude modulation. Suppose that we want to send the binary digital message 11010011. If this message were encoded using on-off keying, the pattern of energy flow values versus time produced would look like:

We can view this as a function of time, d(t). At any time when we are transmitting a 1 from the binary message, the value of d(t) is 1. Whenever we are transmitting a 0, d(t) = 0. Now, if we use such a digital signal to control a light source for a fiber optic transmission system, the signal actually flowing through the system will be described by:

$$d(t) \times \sin(f \times 2\pi t)$$

where `f' is the frequency of the light emitted by the light source. Basically, when d(t) is 1, it is clear that $d(t) \times \sin(f \times 2\pi t)$ is just equal to $\sin(f \times 2\pi t)$. So, during these periods the wave produces is just light of the frequency produced by the light source. Since 0 times anything is 0, when d(t) is 0, the value of $d(t) \times \sin(f \times 2\pi t)$ is 0 corresponding to the fact that no signal is transmitted. Accordingly, the transmitted signal, s(t) will look something like figure 4.15 :

Figure 4.15

When we multiplex two signals using different colors of light, the multiplexed signal can be described by accounting for the frequency differences between the two light sources. Suppose that the two digital signals to be transmitted are described by the functions d1(1) and d2(t). Assume we use red and blue light as in our earlier examples and that f1 is the frequency of the red light used and that f2 is the frequency of the blue light used. Then the combined signal will be described by the formula:

$$d1(t) \times \sin(f1 \times 2\pi t) + d2(t) \times \sin(f2 \times 2\pi t)$$

Of course, these mathematical descriptions only apply to amplitude modulation. Other techniques such as frequency modulation involve different mathematical formulations. They are not really more complicated than what we have developed for amplitude modulation, but a mathematical understanding of amplitude modulation will be sufficient for our purposes.

4.8 Analog Modulation and Bandwidth

We can gain additional understanding of real transmitted signals by considering the transmitted signals that would be produced if we were to encode some more examples of unrealistic "data". We begin by considering the case of a very simple repetitive analog signal. Suppose that we wanted to transmit an analog signal that itself took the form of a simple sine wave. The form of a note produced by a flute, for example, comes quite close to a mathematical sine wave.

In this situation, both the audio signal and the carrier wave would be described by functions of the form $\sin(f \times 2\pi t)$. The main difference between the two would be in their frequencies. We will call the frequency of the analog signal f_A and the frequency of the carrier wave f_c. Given these names, the transmitted signal will be described by:

$s(t) = a(t) \times c(t) = \sin(f_A \times 2\pi t) \times \sin(f_c \times 2\pi t)$

If we pick some values for f_A and f_c, we can plot the function s(t) to get a sense of what the transmitted waves would look like in such a situation. For example, if we let $f_c = 11$ and $f_A=1$ the graph of the function s(t) will look like figure 4.16:

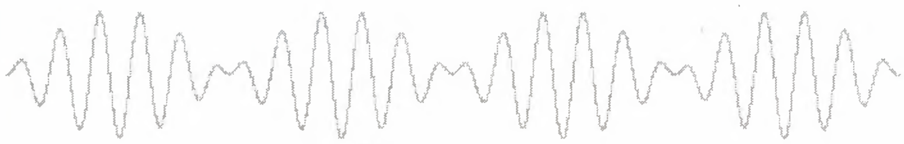


Figure: 4.16

This wave should look familiar. While the match is not exact, it is very similar to the waveform we concluded would be formed when red and green light were mixed together. This is not a coincidence. If you dust off your old trigonometry textbook or still somehow remember a few of the trigonometric identities you learned in high school, you will discover that the "product of sines" identity states that:

$2 \sin z1 \sin z2 = \cos(z1 - z2) - \cos (z1 + z2)$

and that for all angles

$- \cos(z) = \cos(z + \pi)$

So, the function $s(t) = \sin(f_A \times 2\pi t) \times \sin(f_c \times 2\pi t)$ must be equivalent to:

$s(t) = (\cos(2\pi t(f_c - f_A)) + \cos(2\pi t(f_c + f_A) + \pi))/2$

What does a function like $\cos(2\pi t(f_c - f_A))$ look like? We can explore this question by plotting a cosine function. If we plot cos(t) in green and sin(t) in red, the result will look something like figure 4.17:

Figure: 4.17

Basically, a cosine wave is shaped just like a sine wave. It is just shifted over a bit as if it started a bit later than the sine wave. We can also express this mathematically. Another trigonometric identity states that:

$\cos(z) = \sin(z + \pi/2)$

So, we can write

$s(t) = 0.5 \times \sin(2\pi t(f_A - f_c) + \pi/2) + 0.5 \times \sin(2\pi t(f_A + f_c) + 3\pi/2)$

This says that s(t) is just the sum of two sine waves. The first sine wave has frequency $f_A - f_c$ and the second wave has frequency $f_A + f_c$. Both waves are shifted over a bit (one by $\pi/2$ and the other by $3\pi/2$), but this does not change the fact that they are sine waves. So, in general, we see that the result of modulating one sine wave by another is equivalent mathematically to the sum of two sine waves. This explains why the plot of the transmitted signal we produce by modulating a carrier sine wave by another sine wave looks so much like the plot we produced to describe the result of summing the two sine waves that correspond to red and green light. The modulated wave is also equivalent to the sum of two sine waves.

In fact, if we pick the frequencies very carefully, we could create a modulated wave that was mathematically equivalent to a mixture of red and green light. Recall that 460,000,000,000,000 cycles per second is the frequency associated with a shade of red light and that

552,000,000,000,000 cycles per second corresponds to green light. The average of these two values is 506,000,000,000,000. Half of the difference between these two values is 46,000,000,000,000. So suppose we let f_c = 506,000,000,000,000 and f_A=46,000,000,000,000. That is, we use light with a frequency of 506 million million cycles per second as our carrier signal (a light wave of this frequency would be perceived as yellow light), and we encode an analog signal with a frequency of 46,000,000,000,000 cycles per second. The resulting signal,

$s(t) = \sin(f_A \times 2\pi t) \times \sin(f_c \times 2\pi t)$

would, according to the trigonometric identities discussed above, be equivalent to:

$s(t) = 0.5 \times \sin(2\pi t(f_A - f_c) + \pi/2) + 0.5 \times \sin(2\pi t(f_A + f_c) + 3\pi/2)$

which is just

$s(t) = (\sin(2\pi t(460{,}000{,}000{,}000{,}000) + \pi/2) + \sin(2\pi t(552{,}000{,}000{,}000{,}000) + 3\pi/2))/2$

That is, the resulting wave would be equivalant to the sum of two waves whose frequencies are those of red and green light. So, if one were to direct the light produced by a system that modulated a beam of yellow light in this way and passed the light produced through a prism it would separate into distinct beams of read and green.

Figure. 4.18

This is not practically possible. While we can build circuits that act like electronic dimmer switches, we cannot build such a circuit that can vary its output quickly enough to respond to an input signal with a frequency of 46,000,000,000,000 cycles per second. The mathematics, however, are correct and do apply to input signals at lower freqencies.

A simple example of this, beat frequencies, is well known to musicians. Above, we have used our mathematical arguments to claim that if a sine wave is modulated, the result is equivalent to the sum of two sine waves. The mathematics also lead to the conclusion that the opposite must be true. That is, if we add two sine waves together, then the result will look like a sine wave that has been modulated. The closer together the frequencies of the sine waves that are combined, the slower the rate of modulation will be. Musicians use this fact to tune instruments. When two instruments play notes that are almost but not quite the same, the loudness of the sound produced seems to pulsate. As the instruments are adjusted to become closer to being in tune, the rate of the pulsation slows until it stops when the instruments are playing the same note. Even if you don't have any instruments handy, you can still observe this effect. Just find a friend who can hum. If you both start humming the same note and then just one of you slightly changes the pitch at which you are humming you will hear the sound begin to pulsate.

As a realistic example of these phenomena involving information transmission, consider what happens when a radio station, like WABC, broadcasts by modulating the amplitude of a carrier signal. As mentioned above, WABC broadcasts using a carrier wave whose frequency is 770,000 cycles per second. That is, for WABC, f_c = 770,000. The sound waves our ears are capable of hearing range in frequency from about 15 cycles per second to about 20,000 cycles per second. Just as white light is a mixture of light waves of many frequencies, most sounds we hear are mixtures of several frequencies. To keep things simple, however, we will assume that WABC only transmits sounds corresponding to pure sine waves with frequencies within the audible range. So, f_A may range from 15 to 20,000.

The transmitted signal would again be

$s(t) = \sin(f_A \times 2\pi t) \times \sin(f_c \times 2\pi t) = (\sin(2\pi t(f_A - f_c) + \pi/2) + \sin(2\pi t(f_A + f_c) + 3\pi/2))/2$

That is, the signal can be described as the sum of two sine waves with frequencies $f_A + f_c$ and $f_A - f_c$. If the frequency of the sound being transmitted is near the low end of the audible range, say 100 cycles per second, then the radio waves transmitted will be the sum of a sine wave of frequency 770,100 cycles per second and another of 769,900 cycles per second. On the other hand, if the sound wave transmitted has a frequency of 10,000 cycles per second, the radio waves composing the actual signal will have frequencies of 780,000 cycles per second and 760,000 cycles per second.

The important point here is that when the operators of a a radio station like WABC tells you that they broadcast at a frequency of 770 khz, they are lying just a bit. Depending on the audio signal being transmitted, the actual signals emanating from the station's radio tower will be composed of a number of radio waves of frequencies close to but not exactly at 770,000 cycles per second. Assuming that WABC only transmits sound corresponding to audible sine waves, the worst case would be transmitting a sound wave with frequency 20,000 cycles per second. While transmitting such a sound, the actual radio waves produced by WABC would be two sine waves with frequencies of 790,000 and 750,000 cycles per second.

Thus, the radio waves produced by a radio station do not all have precisely the frequency assigned to the station. Instead, they can fall anywhere in a range around the assigned frequency. The assigned frequency is the frequency of the carrier wave that is to be modulated to transmit signals. The range of frequencies actually produced as the carrier is modulated depends on the frequencies of the signals being encoded. The range of audio frequenices being encoded determines the size of the range of radio frequencies actually used. Thus, because audible sound wave have frequencies that range up to about 20,000 cycles per second, the range of radio frequencies actually produced by an AM radio station is about twice this wide. The lowest frequencies used will be about 20,000 cycles per second below the assigned frequency and the highest used will be about 20,000 cycles above the carrier signal. The size of the range of frequencies used in such a transmission system is called the bandwidth of the transmitted signal.

The FCC, which is the US government agency responsible for assigning frequencies to radio stations, had to take this into account when assigning WABC the carrier frequency 770,000. If there were any other AM radio station in the New York City area whose carrier frequency was within 40,000 cycles of 770,000, then the range of frequencies of radio waves produced by this station and those produced by WABC would overlap to some degree. This would make it impossible for you to tune your radio to one station or the other. Your radio would not be able to separate the signals coming from the two radio transmitters.

Similar considerations apply to transmission of digital data through fiber. While we can not practially modulate light waves at frequencies high enough to turn pure yellow light into a form equivalent to a mixture of red and green light waves, even modulation by a lower frequency signal will turn a light carrier of a given frequency into a collection of light waves whose frequencies fall in a range around the carrier frequency. The higher the frequency of the signal that controls the modulation, the wider the range of the frequencies actually transmitted.

Just as the bandwidth associated with radio stations limits how close together the carrier frequencies assigned to AM stations can be, the width of the range of frequencies of light present in a beam transmitted through a fiber will limit how many distinct signals can be multiplexed on the fiber. Recall the imaginary communications system we proposed for Bob and Alice earlier in this chapter. It depends on the ability of a prism to separate the independent signals sent by Bob and Alice so that they can be directed at separate light detectors. In such a system, the bandwidths of frequencies in the modulated signals will actually correspond to the physical width of the beams of light that impact the photoelectric cells. If Bob increases the rate at which 0's and 1's are transmitted, this will increase the rate at which the green light which is his carrier wave is modulated. As a result, the range of frequencies in the beam of green light will increase. Because the rate at which binary data can be transmitted is much lower than the frequency of light, Bob's beam is unlikely to split into beams of different colors when it passes through the prism. It will, however, contain a range of shades of green that will be bent different amounts by the prism. The higher the rate at which the signal is modulated, the wider the beam of shades of green will become. When selecting light frequencies to use as carriers in such a system, the designer must make certain that the frequencies chosen are far enough apart that the modulated beams will not overlap as they leave the prism.

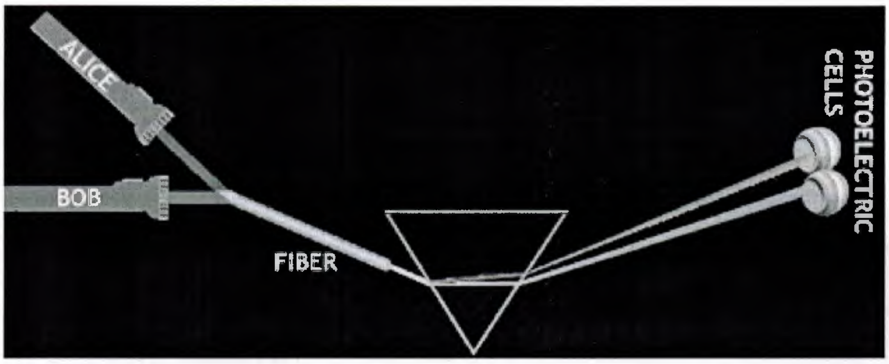


Figure: 4.19

While these claims are true, they don't quite follow from our discussion of analog modulation. The mathematical argument that showed that a modulated signal was equivalent to the sum of two pure sine waves was based on the assumption that the signal controlling the modulation was a pure sine wave. If the modulating signal is a sequence of 0's and 1's, it is about as far from being a pure sine wave as possible. Instead, it is a square wave. To provide a more adequate explanation of the behavior of light waves modulated by a digital signal, we need to explore one more aspect of the mathematical description of waves.

4.9 Digital Modulation and Fourier Series

The conclusions of the preceding sections all seem to hang on what probably appeared to be an obscure mathematical trick. By applying some long-forgotten trigonometric identities, we were able to show that a modulated carrier wave was mathematically (and therefore physically) equivalent to the sum of two sine waves. In fact, however, this mathematical trick is just an example of a very general mathematical and physical properties of waves. All waves are in some sense just the sum of a collection of simple sine waves.

It is perhaps best to first look at the physical justification of this principle, the prism. When we take any sort of light and pass it through a prism, the prism breaks the light up into its constituent parts. The set of colors present in the light leaving the prism is called a spectrum. White light produces a complete spectrum containing shades of all the colors.

Figure: 4.19

Other forms of light that contain only a subset of the possible colors produce incomplete spectra. For example, the spectrum of the light produced when electric current is passed through a tube containing hydrogen (i.e. a hydrogen light rather than a neon light) looks like:

Figure: 4.20

Physically, what the prism is showing us is that any light wave we can produce can be broken up into a collection of simple light waves of pure colors.

While there is no simple equivalent to the prism for sound waves, anyone who has adjusted the bass or treble controls on a car radio should at least aware of the fact that any sound is composed of high frequency components and low frequency components. Many stereos include displays that break down the sound currently being played into many more than two frequency ranges and display the intensity of the sound in each range. Without even leaving the comfort of your device you can see an example of this. The control window produced by running the Macintosh version of

RealPlayyer's streaming media application is shown in figure 4.21.

Figure: 4.21

In the horizontal subwindow below the large "real" icon in the center of the window the program displays a graphical meter showing the volume of the current sound broken down into stereo channels and into frequency ranges within each channel. The subwindow displays a total of 32 vertical bars. The 16 bars on the left describe the left stereo channel, the bars on the right describe the right channel. Within each group of 16, the bar on the left describes the lowest frequency range and each bar corresponds to a higher range of frequencies as one moves to the right. Since the designer's of the program probably didn't expect anyone to use this display as a scientific instrument, they provide no clue exactly what range of frequencies goes with each bar. The height of the pile of brightly colored rectangles in each vertical bar describes the relative loudness of a given range of frequencies within the total sound. Bascially, the bars are an attempt to provide a visual representation of the spectrum of the sound being played.

The mathematical principle corresponding to the physical idea of a spectrum is due to a French mathematician name Jean-Baptiste Fourier. In the early 1800's, Fourier showed that any mathematical function that was "wavelike" (i.e. repetitive and reasonably smooth) was mathematically equivalent to the sum of a collection of simple sine waves. The collection of sine waves whose sum is equivalent to a given function is called that function's Fourier series. These sine waves would all have different frequencies. In addition, they might have different amplitudes and they might be shifted relative to one another, as we saw the two sine waves in our simple sum were shifted by $\pi/2$ and $3\pi/2$. While the only examples we have seen of Fourier series have involved sums of just two sine waves, Fourier's result allowed for the possibility that an infinite collection of sine waves might need to be added together to produce a sum equivalent to a given non-simple wave.

In the last section, we applied the ability to view a wave as the sum of two simple sine waves to the transmitted signal. Both the carrier and the encoded signal were assumed to be simple sine waves and it was in understanding the complex modulated signal that was actually transmitted that we used a simple special case of a sum of sine waves to describe a complex wave. Now, what we want to be able to do is understand the application of amplitude modulation in cases where the encoded signal is not a simple sine wave. Untimately, we would like to be able to understand cases where the encoded signal is a square wave encoding binary data. To start, however, let's think about the implications of Fourier series when applied to the encoding of relatively simple audio signals.

By relatively simple, we mean audio signals that are not sine waves themselves but are equivalent to the sum of a small number of simple sine waves. The most simple such case would be a sum of two sine waves. So, suppose we are using amplitude modulation to encode an audio signal

$a(t) = \sin(2\pi t \times f_A) + \sin(2\pi t \times f_B)$

A plot of what a(t) would look like in the case that $f_A = 2$ and $f_B = 5$ is shown in figure 4.22 below:

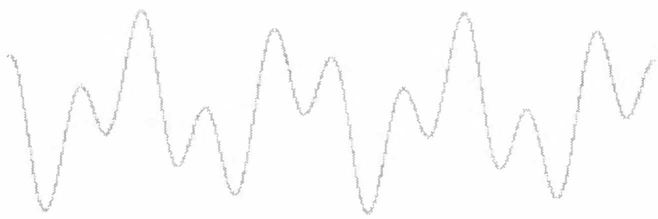


Figure: 4.22

If we transmit this audio signal using amplitude modulation and a carrier wave with frequency f_c, the transmitted signal will be:

$s(t) = a(t) \times c(t) = [\ \sin(2\pi t \times f_A) + \sin(2\pi t \times f_B)\ ] \times \sin(2\pi t \times f_c)$

By distributing the term for the carrier wave, this can be rewritten as:

$s(t) = \sin(2\pi t \times f_A) \times \sin(2\pi t \times f_c) + \sin(2\pi t \times f_B) \times \sin(2\pi t \times f_c)$

Thus, s(t) is the sum of two products of simple sine waves. Luckily, we already know how to rewrite products of simple sine waves as sums of sine waves from the last section. Applying this transformation to each product in the sum above, we obtain:

$s(t) = 0.5 \times [\ \sin(2\pi t \times (f_c + f_A)) + \sin(2\pi t \times (f_c - f_A)) + \sin(2\pi t \times (f_c + f_B)) + \sin(2\pi t \times (f_c - f_B))\ ]$

In the case of an audio signal described by a simple sine wave, we saw that the transmitted signal consisted of a pair of sine waves whose frequencies fell above and below the frequency of the carrier wave by an amount equal to the frequency of the audio signal. Here, we find that the transmitted signal consists of two pairs of sine waves. There is one pair for each of the sine waves that compose the audio signal. Each pair consists of sine waves whose frequencies fall above and below the carrier wave by an amount equal to the frequency of the component of the Fourier series to which it corresponds.

If we modify the amplitude of one of the sine waves in the sum, this change in amplitude is passed on to both members of the corresponding pair of sine waves in the modulated signal. That is, if

$a(t) = K \times \sin(2\pi t \times f_A) + \sin(2\pi t \times f_B)$

then

$s(t) = [K \times \sin(2\pi t \times f_A) + \sin(2\pi t \times f_B)] \times \sin(2\pi t \times f_c)$

$= 0.5 \times [K \times \sin(2\pi t \times (f_c + f_A)) + K \times \sin(2\pi t \times (f_c - f_A)) + \sin(2\pi t \times (f_c + f_B)) + \sin(2\pi t \times (f_c - f_B))]$

It should also be clear that we could repeat this analysis for audio signals formed by adding together more than two sine waves. No matter how many sine waves are added together, the result of modulating a carrier by their sum will produce the same result. For each sine wave in the sum there will be a pair of sine waves in the transmitted signal. The frequencies of the members of each such pair of sine waves will fall above and below the frequency of the carrier signal by amounts equal to the frequency of the corresponding sine wave from the original signal. Thus, if we can find the Fourier series for any audio or digital signal we plan to encode using amplitude modulation, then we can predict the form of the transmitted signal. For each sine wave in the Fourier series of the audio or digital signal, the transmitted signal will include a pair of sine waves with appropriate frequencies and amplitudes.

This is enough to enable us to get a more precise idea of the nature of the radio waves produced when a real audio signal is transmitted. As the bars in the RealPlayer display presented as an example above suggest, the Fourier series for sounds such as voice or music contain a large number of sine waves of frequencies varying from 15 to 20,000 cycles per second. As a result, the signal that would actually be transmitted by a station like WABC would contain a large number of pairs of sine waves centered around the frequency of the radio station's carrier wave (770,000 in the case of WABC) extending no more than 20,000 cycles per second above or below that carrier frequency. If we could visually perceive the spectrum of such a radio signal it would be symmetric around the carrier because each component of the Fourier series of the sound being transmitted results in a pair of similar sine waves above and below the carrier frequency.

Now, we are ready to attack what is really our ultimate goal. What will the transmitted signal look like if amplitude modulation is used to encode a digital signal for transmission. The answer to this question will enable us to describe the actual characteristics of the light waves produced when we signal through a fiber optic cable by rapidly turn some light source on and off to encode 1's and 0's. It will also tell us a good bit about another familiar use of modulation to encode digital information. Those who have used or still use modems to connect to an Internet service provider are also using modulation to encode digital signals. The name "modem" comes from a longer, original name for the device, modulator-demodulator. A modem takes a digital signal and uses it to modulate a sound wave producing a signal suitable for transmission through the phone system.

There are many possible sequences of 0's and 1's that might be transmitted through a fiber or using a modem. Each such sequence will correspond to a waveform composed of rectangular segments. Although such square wave forms are far less smooth than most of the waves we have been considering, Fourier's result still says that they can be described by a sum of sine waves. A different set of sine waves, however, will be required for each different sequence of 0's and 1's. So, the best we can do is get a general sense of what such a sequence might look like. We will pick a particularly simple sequence for our example: a never-ending sequences of alternating 0's and 1's ---010101010101010.... A graph of the signal we have in mind looks like:

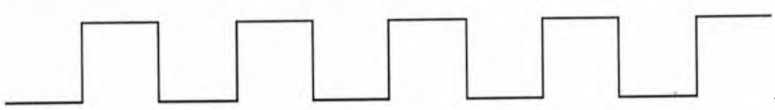


Figure: 4.23

At first, this graph probably does not look like the sum of a collection of sine waves. Fourier's work, however, says that it is. Unfortunately, the collection of sine waves that have to be added together to match this square wave exactly is infinite. One can, however, get a good sense of how the sum works by looking at a small subset of the terms included in the infinite sum.

Consider the shape of the graph of the sum five sine waves shown below.

$$\left[\sin f\pi x + \frac{\sin 3\pi f x}{3} + \frac{\sin 5\pi f x}{5} + \frac{\sin 7\pi f x}{7} + \frac{\sin 9\pi f x}{9}\right]$$

Figure: 4.24

While this isn't a perfect match for our square wave, it is probably closer than you might have guessed we could come by adding just five sine waves together. The terms in this sum follow a clear sequence. Each term is of the form:

$$\frac{\sin n\pi f x}{n}$$

where n is an odd number and f is the frequency of the original square wave. If we continue adding together terms of this form up to the point where n = 19, the resulting function looks like figure 4.25:

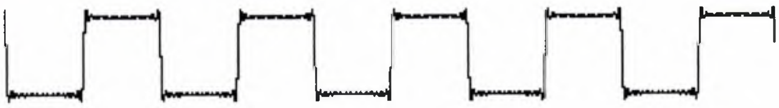


Figure: 4.25

It still does not exactly match the square wave. If you zoom in on the top half of one of the peaks you can still see the wiggles in figure 4.26:

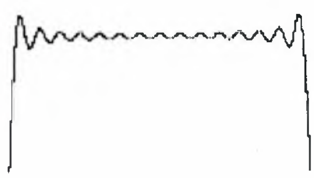


Figure: 4.26

but it is getting quite close. In fact, Mr. Fourier's results show that if you add the sine waves of the form $\frac{\sin n\pi f x}{n}$ together for all odd numbers n, the shape of the result will match the square wave exactly. That is, if we use d(t) as a name for the digital signal, then

Now that we know the Fourier series for this square wave, we can consider the form of the transmitted signal produced by using the square wave to modulate a sine wave used as a carrier. We will call the frequency of the digital signal f and, as ususal, the frequency of the carrier will be named f_c.

$s(t) = \sin 2\pi x f_c \times d(t) = \sin 2\pi f C \times \left[\sin f\pi x + \frac{\sin 3\pi f x}{3} + \frac{\sin 5\pi f x}{5} + \frac{\sin 7\pi f x}{7} + \frac{\sin 9\pi f x}{9} + \ldots\right]$

As we did when working with finite sums of sine waves, we can distribute the multiplication by the carrier wave over the sum. The result will be a sum of terms of the form

$\sin 2\pi x f_c \times \frac{\sin n\pi f x}{n}$

Each of these terms can be rewritten as a sum of two sine waves, namely:

$\frac{1}{2n}\left[\sin(2\pi x(f_c + \frac{nf}{2})) + \sin(2\pi x(f_c - \frac{nf}{2}))\right]$

The actual transmitted signal is the sum of all of these pairs of sine waves. It is the sum of an infinite number of such pairs. Therefore, its spectrum include waves of an infinite number of different frequencies. As with the simpler cases we considered earlier, the frequencies of these waves are centered about the frequency of the carrier wave. The frequencies present are all of the form f_c + nf/2 or f_c - nf/2 for all odd values of n.

When we discussed the frequencies present in the signals transmitted by an AM radio station, we introduced the notion of bandwidth, the size of the range of frequencies present in a signal. In the case of AM radio transmission, the range of frequencies used is limited by the range of frequencies present in audible sound. The differences between the frequencies of the sine waves that compose a modulated signal and the carrier frequency are determined by the frequencies of the sine wave that compose the original signal. In the case of audible sound, there is an upper bound on the range of frequency components present in the original signal. No audible sound has a frequency above 20,000 cycles per second. Thus, the bandwidth of an AM radio station is roughly 40,000 cycles per second.

There is no bound on the frequencies present in the Fourier series of the simple digital signal we have considered. Waves with the frequency f_c + nf/2 will be present for arbitrarily large values of n. In some sense, therefore, the bandwidth of a signal produced by using a digital signal to modulate a carrier wave is infinite. Recall that in the case of AM radio stations, the bandwidth determines how widely spaced the carrier frequencies assigned to stations have to be to avoid having components of signals from different stations overlap. In a situation where modulation is used to encode a digital signal, components from distinct signals will overlap no matter how widely the carrier frequencies are spaced. In theory, this would make frequency division multiplexing of such signals impossible.

The good news, is that for all practical purposes, the signal produced by modulating a carrier wave using a digital signal has a limited bandwidth. If one examines the terms that describe the sine waves that compose the Fourier series of modulated signal representing an alternating sequence of 0'a and 1's, you can see that the terms contribute different amounts to the total signal. Each term is of the form:

$$\frac{\sin n\pi f x}{n}$$

The wave described by the first term has amplitude 2 since its value varies from a high of 1 to a low of -1. On the other hand, the term for n=99 has an amplitude of 2/99ths. Because each term is divided by the value of n, the range of values for the n = 99 term is from 1/99th to -1/99th. The amplitude of an electromagnetic wave or a sound wave is related to its power. Therefore, the n =99 term contributes relatively little power compared to the earlier terms.

You can see this by looking back at our diagrams of waves described by the first few terms of the complete Fourier series for a square wave. The diagrams are already quite close to being square waves. That implies that adding in the values of the waves of higher frequency components of the Fourier series has very little effect on the overall shape of the wave received. The receiver will have to be flexible enough to interpret incoming signals that are not perfect square waves because the transmitted signal will be distorted by noise and other factors. If the original signal is close to but not quite a sine wave the receiver is unlikely to even notice the difference.

We can take advantage of these observations and reduce the bandwidth of the transmitted signal by removing all but the first few components of the Fourier series of the pure square wave before using it to modulate the carrier wave. Removing frequencies is not difficult. It is called filtering. You can do it on your car radio by simply turning the treble knob to the lowest possible setting. This removes the high frequency components from the spectrum of the sound produced by your radios. Sophisticated audio equalizer provide even finer control over frequency filtering. So, to reduce the bandwitdh required to transmit a digital signal using modulation, we simply filter out the high frequency components. The resulting bandwidth will still be large compared to that required for an analog signal. As we saw when discussing AM radio transmission, the bandwidth required for an analog signal is roughly twice the highest frequency component of the signal. To ensure that the signal transmitted is sufficiently accurate, however, when working with a digital signal we cannot filter out any of the first few components of its Fourier series. If we keep just the first five terms, we will have a highest frequency 9 times the bit rate of the signal requiring a bandwidth of 18 times the bit rate of the signal.

4.10 AM and Stereo Transmission

The introduction of stereophonic transmission to AM broadcasting has allowed it to become more competitive with FM stereo broadcasting. However, just as the AM stereo transmission process is very much different than that of FM stereo, so is the requirement for proper audio processing of AM stereo. Special processing requirements are needed in order to maximize both good monaural compatibility and high quality stereophonic transmission simultaneously.

With more AM radio stations converting to stereo broadcasting, numerous transmission problems unique to AM stereo face them. One of the more important concerns is the proper understanding and choice of the audio processing technology for it. As different as AM stereo technology is from existing FM stereo, so are the problems of its audio processing. Because AM stereo is achieved by separate monophonic and difference information modulation paths, FM type left and right audio processing does not provide the best results for AM stereo. Under varying amounts of separation, left and right processing can cause up to a theoretical 6 db loss of monophonic loudness which can result in the loss of coverage of a radio station's existing fringe areas. The following is a brief discussion of these problems and their cures.

4.11 FM-Type Left and Right Stereo Limiting

In FM stereo transmission, the left and right channel information can be fundamentally described as sent via the same transmission path as shown in figure 1. The following equation describes the general stereo transmission signal:

$$f(t) = (L+R) + (L-R)\sin wt + P\sin (w/2)t$$

In the above equation, wt represents the 38 Khz subcarrier and Psin (w/2)t represents the fixed amplitude 19 Khz pilot. From inspection of the equation, it is apparent that when L and R are equal to each other in phase and are limited to a peak amplitude of 1, the (L-R) term equals 0 and the peak amplitude of the output function f(t) equals 2 plus the pilot amplitude. Upon inspection of the equation it is also found that the peak amplitude of f(t) is still equal to 2 plus the pilot amplitude when left only or right only single channel conditions are transmitted. This occurs because the equation's sin wt function reaches numerical extremes of 1 and -1 and forces f(t) to equal either (2 x L) or (2 x R) during those conditions.

Thus, when levels are properly matched, separate (discrete) left and right limiting of audio signals will result in the identical 100% modulation limits for left channel only, right channel only, or both together (monaural during stereo) transmissions. This demonstrates that separate left and right audio processing is the most appropriate choice for maintaining modulation limits of FM stereophonic type transmissions.

4.12 AM Stereo Limiting Requirements

AM Stereo broadcasting has brought about a need for a different form of stereo audio limiting. This is because the left and right transmitted audio channels are first transformed into L+R and L-R terms through a matrix summer and subtractor and it is THESE components that are applied to two modulation points and transmitted as shown in figure 2. The basic AM and Motorola C-QUAM® equation for a transmitted wave is as follows:

$f(t) = A[1+M(L+R)] \cos(wt + Z)$

In the basic equation, either AM monaural audio or the AM stereo L+R term is used to vary the A[1+M(L+R)] value (modulated amplitude) of the f(t) equation. In stereo, L-R plus the other terms which make Motorola C-QUAM® are used to vary the Z value (modulated phase) of the f(t) equation.

Because the algebraic sum and difference of left and right channels occur PRIOR to the points of the modulation, AM stereo broadcasting is best supported by stereo "matrix L+R/L-R" limiting. This enables the final limiting action to be shifted to the sum (L+R) and difference (L-R) axis of the stereo sound field where the actual transmitted components exist. This method produces significantly improved performance over the discrete audio processing types used for FM which operate only on the left and right channel axis.

Displaying AM Stereo Limiting Patterns

When monitoring the X-Y lissajous pattern produced by the right and left outputs of the limiters they can be easily seen on the oscilloscope. This pattern can also be seen complete with any distortions which may be caused by the transmitter by monitoring from the left and right audio outputs of the radio station's high quality stereo modulation monitor. If the limiters have L+R and L-R outputs instead, the patterns at these outputs will be shifted counter clockwise by 45 degrees from those illustrated. Field experience has shown that once familiarity with these patterns is gained, they are often more helpful in checking for proper processing alignment and show more information about what is being transmitted than any other form of modulation monitoring system.

Oscilloscope X-Y display of the right and left limiter outputs of conventional stereo limiting. When applied to AM stereo transmissions, the amplitude limit levels of the left and right channels must be set equal to each other for proper stereo balancing. As shown, the levels are perpendicular to the right and left channel axis and intersect with each other to form the L+R and L-R modulation limits. The L+R axis represents the main monaural component transmitted by the AM envelope of the transmitter and the L-R axis represents the main stereo information component transmitted by the phase modulation of the carrier frequency. As long as the program input is mostly monaural, this limiting system produces nearly full 100% envelope modulation and monaural reception remains normal.

However, such limiting creates serious monaural transmission and reception problems during varying stereo conditions. When stereo inputs temporarily shift to the full left only (vertical) or right only (horizontal) modulations, stereo reception is acceptable but monaural is not. The L+R modulation component is forced to drop to 50% as is shown by the dotted line intersection of the lower right modulation scale with the tips of the left channel or right channel limit levels. This indicates an immediate 6 db drop in loudness in monaural reception. This is obviously an unacceptable condition to AM broadcasters who are concerned about monaural listeners since the existing monaural coverage and loudness is reduced.

Although most stereo program material does not contain significant amounts of single channel passages, this form of limiting can cause significant losses of monaural loudness and coverage on nearly any stereo program material. The losses are usually directly proportional to the stereo content and become greater as stereo separation increases.

4.13 Basic Stereo Matrix L+R/L-R Limiting

The oscilloscope X-Y display of the right and left limiter outputs of full monaural support matrix limiting. With this system, the output levels of the L+R and L-R are adjusted for equal modulation levels which is also the point of maximum separation. As shown, the amplitude limit levels are perpendicular to the L+R and L-R axis and intersect with each other at the left channel and right channel axis. When stereo inputs temporarily shift to the full left only or the right only axis, these limit levels allow the L+R component to remain at a 100% modulation which maintains full monaural reception compatibility during such transmissions. The shaded area shown in the illustration shows the increased areas of monaural support modulation produced by this system as compared to the figure 3 conventional left and right limiting.

Unfortunately, further analysis shows that this will have a 6 db INCREASE in the single channel stereo receptions. While this obviously is going to be noticeable to listeners, critical listening tests have demonstrated this to be a far more acceptable phenomena than the "instant" LOSS of 6 db in loudness in monaural reception. It should also be kept in mind that the majority of stereo program contents do not contain full single channel transmissions.

4.14 Modified Stereo Matrix L+R/L-R Limiting

Under light and moderate amounts of limiting, full matrix processing produces outstanding results in both monaural and stereo. Heavy amounts of limiting or processing can produce different results. Heavy or extreme levels of audio processing as demanded by many existing AM radio stations may cause certain types of overloads in present stereo decoding and reception techniques. In an effort to reduce the chances of these problems, a modified full matrix processing has been developed by Circuit Research Labs, Inc.

The oscilloscope X-Y display of the right and left limiter outputs of the CRL modified monaural support matrix limiting system. The significant difference between this limiting pattern and the one shown in figure 4 is visible in the left and bottom corners of the pattern. Here, the negative going corners formed by the L+R and L-R axis are removed by adjustable single channel left only and right only limiting networks. This system allows full monaural compatibility during the majority of stereo conditions, but causes a reduction of L-R and negative peak L+R modulation levels during extreme left only or right only stereo conditions. In the illustration, the single channel limits are shown set for a left or right only L+R negative limit of 70% instead of the 100% level which would occur without such limiting.

Modified matrix system is designed to reduce the potential problem areas (the negative single channel corners) associated with stereo transmissions. At the removed corners shown in the figure, both L+R and L-R modulations are at maximum and can cause decoding difficulties.

If high density negative peak L+R modulations are allowed to consistently reduce the transmitter carrier, the L-R decoding process has little or no carrier to demodulate. The result can be that either stereo decoding returns to monaural or produces local decoder clipping distortions caused internally in the decoder. Depending upon the degree of processing used and maximum L+R modulation depth, the single channel limiting network can be adjusted from a point having only minor effects (for stations employing small amounts of overall processing) to a level which prevents or substantially reduces stereo receiving problems (when heavy monaural support processing levels are employed).

4.15 A Modified Matrix Limiter

The basic structure of a modified matrix limiting system is the basis for the CRL SMP-950 Stereo Matrix Processor which is the main unit for CRL's AM Stereo Audio Processing line. The block diagram shows the minimum elements needed to produce both compatible AM mono and high quality AM stereophonic transmissions. The basic elements include mono support compression, NRSC pre-emphasis, multi-band pre-emphasis limiting, final L+R/L-R limiting, NRSC low-pass bandwidth filtering and low distortion single channel negative peak limiting which is adjustable for 70% to 75% negative peaks during left only or right only conditions.

CHAPTER V:
PROPOSAL FOR A DIGITAL BROADCASTING SYSTEM IN AM BANDS

Digital transmission methods offer interesting advantages, especially in frequency ranges which, until today, have been used differently. These advantages are mostly only of a technical character (in the sense of optimization of quality or a better use of the spectrum) and only represent a real improvement for the customer if they can be introduced with reasonable cost and effort. This is especially true for AM in general, and particularly for international short-wave broadcasting, because of the widespread use of these media world-wide.

The most important characteristics of AM broadcasting in the three classical frequency bands are the following:

Frequency-ranges
148.5-283.5 Khz (Long Waves, LW)
526.5-1606.5(1705) KHz (Medium Waves, MW)
3.2-26.1 MHz (Short Waves, SW)

Channel-spacing
10 (9) kHz for Double Side Band (DSB)
5 (4,5) kHz for Single Side Band (SSB)

Transmission characteristics

- reasonably stable propagation in LW and MW during day time (ground wave propagation),
- essentially ionospheric propagation with multiple reflections, Doppler effect and Doppler spreading in MW during night time and SW day and night time,

- presence of interference and jamming, especially in the SW frequency bands, due to congestion in the frequency spectrum (multiple re-use of channels, receiver non linearities and/or poor filtering characteristics, etc...).

This results in the following technical requirements for an efficient digital broadcasting system:

- improved reception quality (immunity to fading, to interference, constant quality, bandwidth utilization, optional stereo reception),
- compatibility with existing AM services (channel-spacing, co-channel and adjacent channel interference),
- necessity during a long transition time to continue to use already installed transmitters and antennas with minimal modifications,
- no significant increase in power consumption or peak power transmission on the transmitting side,
- compatibility with low cost and low consumption AM receivers and available technology,
- no increase, and if possible a decrease, in receiver operation complexity.

This, of course, needs a flexible system allowing a progressive introduction capability for the transition period as well as full compatibility with future full digital programmes.

The simplest solution for this transition consists in a compromise where the same transmitter and the same channel are used to broadcast simultaneously analogue and digital programmes (Simulcast).

As a consequence, this leads to an improved service to the listener, as well as new services

- improvement of reception quality (immunity to fading, bandwidth occupancy, optional stereo reception),

- improvement in receiver operation (Listener friendly receiver) programme searching, and, especially for short waves, automatic frequency changing,

- introduction of new data services Programme Associated Data (PAD), or accessibility to radically new independent services (document transmission, fixed images, personal data transmission, ...)

5.1 Types of modulation and selection criteria

5.1.1 Bit rate to transmit

The first consideration in the design of a digital audio broadcasting system is the minimum bit rate which has to be transmitted in order to achieve a perceptual quality at least as good as - and if possible better than - the present quality in the considered bandwidth (LW, MW, SW).

Many existing coders already provide a good quality at bit rates between typically 12 and 48 kb/s. The difference between the bit rates lies essentially in the transmitted audio bandwidth and the capability for stereo transmission.

5.1.2 Channel coding

Due to the poor quality and the instability of the transmission link, especially in the presence of fading and interference, some form of channel coding is mandatory to ensure integrity of the received binary stream.

The best known channel coding systems use high diversity convolutional encoding (known as Trellis Coded Modulation, TCM) associated with interleaving, the purpose of which is to increase robustness to short term reductions of Signal to Noise Ratio (SNR).

5.1.2 Modulation scheme

There are mainly two different modulation schemes which are known as serial modulation and parallel modulation. Both have to convey a high bit rate in a very narrow channel: this needs a very efficient modulation system in terms of the ratio b/s / Hz. This eliminates simple modulations schemes like for example QPSK (Quadriphase Phase Shift keying), DQPSK, etc... which have a poor efficiency.

To comply with this constraint, multi-amplitude/multiphase systems have to be used, and more generally multi-level complex modulations known as QAM (Quadrature Amplitude Modulation), APSK (Amplitude/Phase Shift keying), etc... which can provide a much higher efficiency.

The first modulation scheme is serial (single carrier modem) a complex symbol is transmitted on the carrier at each signalling interval.

In the presence of multipath effects, the received signal is a mixture of many different successive symbols: the receiver has to separate them before demodulation, by means of an equalizer.

The complexity of the equalizer is proportional to the square of the delay spreading. When the channel is unstable (especially in the SW band), this complexity is increased by the fact that one has to accurately estimate a rapidly changing multipath configuration. In addition, the implementation of the equalizer (Decision Feedback Equalizer, DFE) suffers from the problem known as error propagation which arises when the binary output stream is not perfectly decoded and/or during deep fading.

The second modulation scheme is parallel (multiple carrier modem, similar to DAB Eureka 147/COFDM): many different subcarriers are independently modulated by complex symbols at a reduced symbol rate. The effect of propagation lies essentially in the frequency domain: the complex gain of the channel is a function of the exact frequency of the considered individual carrier, and has to be estimated in order to demodulate this carrier.

Demodulation is currently done using Fast Fourier Transforms (FFT) associated with time and frequency filtering/interpolation to estimate gains.

In this scheme, the simultaneous reception of different delayed versions of a given symbol leads to the use of a guard interval corresponding to the maximum delay spreading. Consequently, the useful symbol duration is slightly less than the inverse of the symbol rate.

5.1.3 Peak to RMS ratio of the transmitted waveform

For a given peak transmission power, the effective average transmitted power depends upon the modulation scheme.

It is commonly admitted that the crest factor of a serial modem is 4-5 dB with a Nyquist filter (alpha = 0.25) i.e. that for a 100 kW transmitter, the mean transmitted power is between 30 and 40 kW.

Concerning a parallel modem with a sufficient number of sub-carriers, the crest factor is 10-12 dB (in theory, it is equal to the number of carriers, but 10-12 dB guarantees that the distorsion induced by saturation is negligeable). The same 100 KW transmitter should transmit only 10 KW or less in these conditions. In practice, field tests show that the crest factor can be reduced to 4-5 dB without significant degradation of the performances: the effect of increased distorsion is roughly equivalent to a slight SNR reduction. The crest factor criterion is no longer a decisive advantage of the serial modem over the parallel modem as the practical value for the latter can be drastically reduced to the same level as the serial modem.

In addition, the measured crest factors of the serial modem greatly depend upon the exact signal constellation used. They are greater for complex multilevel modulations than for the simple phase modulation techniques in use today. We can conclude therefor that modulation scheme has no effect on the crest factor of the parallel modem.

5.2 Waveform bandwidth

For the parallel modem, the frequency spectrum of the signal is quasi-rectangular which allows a better channel occupancy with good adjacent channel protection as already demonstrated with similar systems.

For the serial modem the channel occupancy has to be reduced in order to take into account the practical Nyquist filter shape which cannot offer the same adjacent channel protection.

This results in a reduced bitrate for the serial modem with a given channel spacing.

5.3 Required computing power at the receiver side

Considering two equivalent digital broadcasting systems (same bitrate, same bandwidth, same channel coding efficiency) the difference between the serial and parallel modem lies essentially in the computing power devoted to synchronization and demodulation.

In the presence of variable propagation conditions, the computing power is constant and comparatively low for the parallel modem.

Whereas it is basically higher for the serial modem and it rapidly increases with the delay spread of the propagation channel. In severe propagation conditions (i.e. short waves), the required computing power can even exceed the capacity of today's or the near future's standard low cost/low consumption technologies.

5.4 Necessity of a single solution for all AM frequency bands (LW, MW, SW)

If different systems are proposed for different frequency ranges, this leads to more complex solutions for the receivers. Therefore, it will be more expensive and more power-consuming.

Consequently a single solution is highly preferable, especially if it can fit any situation (available bandwidth, required bitrate, robustness to different propagation/jamming conditions) without any modification of the receiver.

5.5 Proposed Solution

The digital system proposed by Thomcast, is the result of an optimization between possible data rates, bandwidth, channel coding, complexity, quality and flexibility.

Skywave 2000 provides a single solution for all AM frequency bands (LW, MW, SW) which will bring benefits to the listener and to the radiobroadcaster in terms of simplicity of receivers, economies of scale, and a wider introduction of the new digital transmission mode. It is the result of a global system approach considering both existing receiver and transmitter techniques and easily implemented at low cost with the technology available today.

In addition, Skywave 2000 approaches the problem of the transition period between the introduction of digital AM Radiobroadcasting today and the future fully digital multimedia service by offering a progressive, compatible (digital and analog) signal which can be received by both todays conventional consumer receivers and by future low cost digital receivers.

Skywave 2000 is based upon a parallel modem which has been proved to be an efficient, reliable and flexible technical solution.

Its incremental architecture allows easy and transparent adaptation to bandwidth, bitrate and level of protection which are required in all present and future implementations, without any change on the receiver side.

The transmitted signal consists of:

- one kernel group of carriers of 3 kHz total bandwidth containing all the signals which are necessary for frequency synchronization, time synchronization, remote control of the receiver, and transmission of a basic bit stream of 8 kb/s with 64 QAM in normal mode and 6 kb/s with 16 QAM in fall back mode.

- a number of additional groups of carriers, each of 1.5 kHz bandwidth, conveying a nominal bitstream of 4 kb/s of audio together with a little more than 200 b/s of data ; the number of additional blocks depends upon the total available bandwidth.

This allows signal bandwidths of 3 kHz, 4.5 kHz, 9 kHz... with bitrates of 8 kb/s, 12 kb/s, 24 kb/s ...(64QAM)

5.5.1 Signal formatting

The basic frame length is 18 ms, corresponding to a useful symbol duration of 15 ms with a guard interval of 3 ms.

The sub-carriers are at multiples of 66.666 Hz (1/15 ms). This spacing has been chosen taking into account the maximum propagation delay spread.

The frames are grouped into bursts of 16 frames (288 ms) labeled 0...15.

In the kernel group of sub-carriers (47 sub-carriers), three are transmitted unmodulated: they are frequency references for fast acquisition and doppler tracking. In the kernel group, all the carriers of frame 0 are unmodulated, and have relative phases ensuring a very low peak factor: this is the time synchronization waveform.

In every group (kernel group, and additional groups of 22 carriers), frames 0, 4, 8 and 12 contain gain references (unmodulated symbols) on even sub-carriers (frames 0 and

8) or odd sub-carriers (frames 4 and 12). This represents an average of 1 reference symbol for each group of 8 symbols.

In each frame, carriers which are neither frequency references, nor time synchronization nor gain references are symbols conveying audio or data. These symbols are modulated using TCM with 64 QAM at 4 bits/symbol (nominal), 16 QAM at 3 bits/symbol (fall back bit rate) or 256 QAM at 6 bits/symbol (maximum bit rate).

In order to improve the channel estimation, and taking into account their low proportion (1/8) reference signals are transmitted at a substantially higher level than free symbols, typically 3 to 6 dB.

5.5.2 Audio source coding

The audio source coder for this type of application is mainly defined by the available bit rate which is inherently variable according to the available bandwidth and to the degree of protection.

A minimum of 6 kb/s (reduced bandwidth, SSB-like speech only) to a maximum of 48 kb/s (stereo music) can be considered like the proper limits of an efficient and convenient system.

Consequently, the chosen coder will be the one which provides the best perceptual quality in real transmission conditions, i.e. taking into account the effect of transmission errors. It will have to be slightly modified - and its useful bit rate reduced - to add protection bits to critical parameters (gains...) in order to ensure graceful degradation of the quality in bad transmission conditions.

5.5.3 Other data

A highly protected low bit rate data stream is devoted to the remote control of the receiver (internal service data). It conveys the transmission parameters: modulation format (16 QAM, 64 QAM, ...), interleaving depth, total bandwidth, ...

A second bit stream, proportional to the total occupied bandwidth, transmits data directly related to the actual transmission: textual informations, next frequency to use, service data, etc...

Finally, in some circumstances, a portion or the totality of the bit stream conveying audio data can be temporarily replaced by bursts of other data (fixed images...): the audio coder bit rate has to be locally reduced in this case, ideally without loss of quality (this can be done during silences, as an example).

5.5.4 Compatibility with existing receivers (Simulcast)

In the transition phase, compatibility with existing standard AM receivers can be achieved by simultaneous transmission of:

- a half bit rate version of the digital audio system

- a compatible SSB transmission of the analog signal, with residual carrier and possibly vestigial side-band to improve quality in the presence of fading and to increase adjacent channel protectio

The digital signal has to be on the high frequency side of the audio signal, since standard receivers attenuate strongly high frequencies the digital part of the transmission is then heard as a weak high frequency unstructured noise.

In addition, if correctly filtered and amplified at the transmitter side, the analog part of the signal has no effect on the digital part.

5.6 Integrated solution

The Skywave 2000 program is conducted with a global system approach from digital source and channel coder at the transmitting side up to digital demodulator/decoder at the receiving side.

The final objective is to develop a component which can be inserted both in the existing radio sets and into the new digital ones.

Features of this component will provide the usual functions of AM receivers and will integrate the new digital demodulator/decoder.

Beyond the sound quality improvements, the digital capabilities integrated into the component will allow new services within the receiver. The component will be very low cost and very low consumption (less than ten milliwatts) and should preferably be able to work with 1.5 V batteries. The component will integrate a 100 MOPS DSP core (Digital Signal Processor), a microcontroller core, several analog converters and a set of on-chip peripheral modules.

The propagation model is an augmented Watterson model with a maximum of 3 different paths. Each path has a deterministic (fixed amplitude) component and a fluctuating (random, Rayleigh) component so that its amplitude has a Rice distribution law. The average power of the two components is adjustable. The frequency offset of the deterministic component is variable. The Doppler spread of the random component can also be modified. The maximum delay of each path is 8 ms, except for the first one which has a delay of 0 ms (not including processing time).

In addition, one can add up to 3 narrow band jammers, with variable level and frequency.

The simulator has an integrated help file. It can be locked/unlocked (password required).

The user interface is based upon menus and dialog boxes.

The user can store on disk and retreive from disk up to 20 different custom configurations.

- a combined analog / digital low power exciter operating in the SW frequency band,

- a standard short wave consumer receiver dedicated to receive the analog programme,

- a modified short wave consumer receiver dedicated to receive the digital programme. The modification consists mainly in an implementation of an IF2 output with a wider IF2 filter connected to the digital decoder.

The demonstrator developed by Thomcast for experimentation and demonstration allows real AM broadcasting operating conditions and offers maximum of flexibility as concerns transmission channel characteristics (already described) and transmission modes.

Among the large choice of transmission modes:
- standard AM DSB,
- SSB: USB or LSB,
- simulcast (Analog compatible AM + Digital) with two versions:
  . analog programme within USB or LSB

- full digital.

Different configurations are available for simulcast and full digital modes. From a nominal 9 kHz RF bandwidth fully compatible with ITU channel allotment and spacing to an extended 12 kHz RF bandwidth compatible with the audio bandwidth of modern existing transmitters (PDM or PSM). The extended RF bandwidth mode allows to demonstrate, at the current stage of audio compression techniques, the highest reachable audio quality and the possibilities for additional data services that a digital system like Skywave 2000 can offer to AM broadcasting if channel allotment and spacing is revised in the future.

This extended mode leads to:

- a digital stereo audio at a usable bit rate of 32 kb/s and 10 kHz audio bandwidth in normal mode and a usable 24 kb/s in fall back mode,

- additional data services with a usable bit rate of 1250 b/s.

- Skywave 2000 tests results and progress

In less than one year, the system developed by Thomcast has demonstrated real fast improvement capability:

- a simulcast mode where an analog programme was only accessible by a SW receiver equiped with SSB mode in June 1996 and now where it is received by means of any existing SW consumer receiver without noticeable disturbance due to the presence within the same HF channel of a digital programme and / or analog programmes broadcasted within the adjacent channels,

- an audio digital programme which started in June 1996 with a 2.7 kHz bandwidth voice at a 4800 b/s usable bit rate and now has a 10 kHz bandwidth stereo audio programme at a 32 kb/s usable bit rate,

- additional data services limited to 44 b/s in June 1996 and now with a 1250 b/s usable bit rate.

Above described experiments have proven the principal capabilities of Skywave 2000 as a system for digital short-wave broadcasting. Since the requirements of Short Wave in terms of system robustness and propagation channel characteristics can be regarded as more severe than for Long Wave and Medium Wave transmissions, Skywave 2000 will also be perfectly suitable for those lower frequency ranges.

The stage that will be reached by Skywave 2000 at NAB 97 Las Vegas features nearly the final system and can be considered as a good basis to start on-air experiments with valuable comparison with current analog services as concerns:

- audio quality improvement,

- area coverage which will normally require much less power for digital transmission to cover the same area with better signal quality,

- compatibility with the existing services.

Skywave 2000 represents a valid system for all AM frequency ranges. It constitutes the current Thomcast contribution for the two international on-going programmes (Eureka 1559 NADIB and DRM Digital Radio Mondiale). Nevertheless, there are some areas which will have to be investigated in more detail in the future, and eventually also a number of improvements, to be included in Skywave 2000.

The future tasks can be sub-divided into three categories:

- further investigations to improve source and channel coding:

. stepwise quality reduction in case of propagation channel impairment, (graceful degradation)

. automatic elimination of interference and jamming signals,

. improved data compression techniques

- investigations of high power AM transmitters:

. evaluation of the state of the art PDM/PSM transmitters,

. necessary modifications in modulators and RF amplification chain,

. investigations of linearity requirement as function of modulation parameters.

- creation of a receiver standard with low cost, low consumption objectives:

. development of a suitable chip-set.

Today's AM and SW broadcasters and the public have a strong interest in conserving use of the unique characteristics of their propagation media well into the future.

It has been shown that a technical solution for digital modulation technology in today's AM frequency bands exists and can be easily implemented.

The Thomcast Skywave 2000 system offers a progressive strategy for introducing digital modulation:

- in all AM Bands (LW, MW, SW),

- in compatible format (Simulcast of digital and analog programmes usable by old and new receivers),

- with maximum flexibility for evolution from analog, to simulcast, to fully digital as the receiver base evolves.

CHAPTER VI:

Conclusion

During the last few decades only few innovations have been introduced into the broadcasting technology on the AM bands (150 kHz - 30 MHz). Amplitude modulation only requires a simple receiver; although this can be a distinct advantage, it can also reduce the quality of transmission. This is a result of audio poor bandwidth and typical propagation conditions, such as, for example, ionospheric instabilities and disturbance from other transmitters.

The poor transmission quality that is associated with AM transmissions arises mainly from the modulation methods and less so from the frequency band. Where the amplitude modulation is replaced by digital modulation, it will be possible to obtain a good quality transmission, maintaining long-distance propagation at the same time.

Because of its non-sensitivity to echo interferences, suitable digital transmission can avoid quality degradations that are typical of analogue transmission.

Digital channel spacing is chosen to be compatible with conventional AM band channel spacing. Because of this, transition from analogue to digital services in AM bands can be made on a channel-by-channel basis.

Todays AM transmitters are generally compatible with digital modulation inputs. The only modification required concerns the oscillator stage mostly.

An AM transmitter is rapidly and simply converted to provide digital modulation. Since all the AF and RF sections of the transmitter remain substantially unmodified, the broadcaster can easily revert back to AM analogue transmissions if necessary.

Radio propagation in MF and HF is influenced by the ionosphere. This is evident for short waves but also for medium waves especially at night times.

Serial modems have to use equalization techniques if fading exists in the propagation area. This will be the case for short wave and upper medium wave, but not for long wave. The receiver thus may have different modes, depending from the conditions in the channel.

For the application of an equalization, measurements of the characteristics of the channel have to be made. These measurements are made with the aid of a periodical test sequence embedded in the modulation stream. With this information available, adaptive equalization is accomplished.

The digital AM receiver is conventional from RF stage, mixer and IF stage. Automatic gain control is necessary like in AM receivers. The IF is succeeded by an analogue or a digital I/Q down mixer to the base band. From there on digital demodulation, error correction and eventually audio decoding follow.

With the aid of the test sequences the receiver is able to switch automatically from digital to analogue reception. The analogue reception also can be done by a signal processor, leading to optimum parameters for analogue demodulation too.

The new method, called T^2M (Telekom-TELEFUNKEN-Multicast), is a multicast transmission without degradation in the analogue or the digital signal, because there are up to 3 HF channels simultaneously transmitted.

During the transition period, it is necessary to have digital transmissions beside AM transmission, else no digital receivers will be produced. On the other hand, switching to digital transmission instantly all previous listeners would be lost.

The multicast offers the way out of this dilemma without great additional cost at the transmitters side. The only costs will be a new oscillator stage and the necessity to restrict the total modulation degree of the transmitter to a maximum of 100%. In some experiments the modulation degree for the AM channel was 70% and for both digital channels together it was 30%.

The modulator needs an audio bandwidth of more than 15 kHz and a good linearity.

T^2M allows several modes:

- analogue plus 2 digital channels
- analogue plus lower digital channel
- analogue plus upper digital channel
- analogue replaced by digital channel

We have the situation, that the upper or lower adjacent channel to the analogue channel has to be free. But the benefits to all broadcasters are obvious, so a possibility for coordinating should be realistic and in the HF bands the antenna has sufficient bandwidth. For reception of the digital transmission the receiver has merely to be tuned to the digital channel, because the receiver cannot distinguish whether this is a multicast transmission or a separate digital transmission.

The effective night coverages of the digital transmissions are considerably bigger than those of the analogue transmission, despite the significant difference in transmitter powers.

The dynamic range of the received audio signal corresponds exactly to the CD audio source signal, and is far better than that of FM transmissions. Of course, the received audio signals are completely free from all usual disturbances associated with AM analogue transmissions.

REFERENCES

1) Ascend Communications, Pipeline 75 User's Guide: Remote LAN Access ISDN Bandwidth on Demand, Ascend Communications, 1275 Harbor Bay Parkway, Alameda, CA, 94502, USA, Part No. 7820-0280-002, August 7, 1996.

2) Barry Truax, Handbook For Acoustic Ecology, Second Edition, 1999, Originally published by the World Soundscape Project, Simon Fraser University, and ARC Publications, 1978

3) Bennet G., Electricity and Modern Physics, Edward Arnold Publishers, London, 1971

4) Chris Hays, AM Powerside, Society of Broadcast Engineers Chapter 36, Los Angeles, http://www.oldradio.com /current/powrside.zip, depicted on Jan 2001.

5) Jim Lesurf, Amplitude Modulation and demodulation, University of St. Andrews, St Andrews, Fife KY16 9SS, Scotland. http://www.st-and.ac.uk/~www_pa/Scots_ Guide /RadCom/intro.html, depicted on February 2001

6) K. Feher, Advanced Digital Communications: Systems and Signal Processing Techniques, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1987.

7) Lajos Hanzo, Amplitude Modulation (Am) Communication Systems http://www.owlnet.rice.edu/~elec301 /Projects99 /amcomm/

8) P.S.K. Leung and K. Feher, FQPSK --- A Superior Modulation Technique for Mobile and Personal Communications, IEEE Trans. Broadcasting, Vol. 39, No. 2, June 1993, pp. 288-294.

9) Radio Communication Theory 1990-1997, University of California.

10) Radio, Basic Physical Principles, Encyclopedia Britannica at www.Britannica.com , depicted on May 2001.

11) Ronald R. Jones Audio Processing, for AM Stereophonic Transmissions,Circuit Research Labs, Inc. Tempe, Arizona U.S.A. Revised 1/93

12) T^{2}m - The State Of The Art Solution For Migration From Analogue To Digital Broadcasting, Telefunken Sender system. Berlin, http://www.broadcast.be, depicted on February 2001

13) Transmitters and receivers, www.homeworkhelp.com, depicted on March 2001

14) W. Gao, M.A. Soderstrand and K. Feher, Gaussian Filter Screens TDMA and Frequency-Hopping Spread-Spectrum Signals, Microwave and RF, May 1995, pp. 17-19.