The Merits of Modulation Systems & Modes of Transmission

Here is some useful information for those being introduced to the principles of Modulation and the various Modes of Transmission used in Amateur Radio

by Lloyd Butler VK5BR

(Originally Published in Amateur Radio July 1991)

Introduction

0ver the years, we have phased out straight amplitude modulation on the HF bands and replaced it with single sideband. On VHF we give preference to frequency modulation. What are the merits of these various systems and why do we choose one in preference to the other? We might expect the answers to these questions to be well understood by the seasoned amateur radio operator, but perhaps not so well understood by our less experienced or novice operator. This article is essentially written for the latter, but a little bit of revision concerning some of the principles might not go astray for all of us.

In the paragraphs which follow we discuss the fundamental differences between the various modulation forms and compare them with reference to such characteristics as bandwidth, noise immunity and transmitter efficiency. Initial discussion involves the various forms of modulation in relation to speech transmission, and this is followed by their application in encoded modes of transmission such as hand-sent Morse, machine telegraphy and other digital systems. A brief reference is also made to their application in television.

Amplitude Modulation

In amplitude modulation (AM) the amplitude of the radio frequency (RF) carrier wave is varied as a function of the instantaneous voltage of the modulating signal. When the modulating signal goes positive, the carrier wave amplitude is increased. When the modulating signal goes negative, the carrier wave amplitude is decreased. The degree of modulation is expressed as a percentage of maximum modulation possible without distortion of the signal information. Figure 1 shows the carrier wave modulated 100 per cent by a sine wave modulation signal. The carrier wave amplitude is doubled by the most positive going excursion of the modulating signal and the amplitude is reduced to zero by the most negative going excursion of the modulating signal.

The waveform shown in figure 1 is a plot of carrier amplitude on the Y axis as a function of time on the X axis, and this is often defined as being plotted in the time domain. lf we plot the amplitude on the Y axis as a function of frequency on the X axis, often defined as being in the frequency domain, we get a different picture. Figure 2 shows, in the frequency domain, a 1MHz carrier frequency modulated 100 per cent by a 1000Hz sine wave modulating signal. The carrier frequency at the centre is the same amplitude as if it were unmodulated. However, there are two side frequencies created, one equal in frequency to the carrier frequency plus the modulating frequency, and one equal in frequency to the carrier frequency minus the modulating frequency. The amplitude of each of the two side frequencies is half that of the carrier frequency.

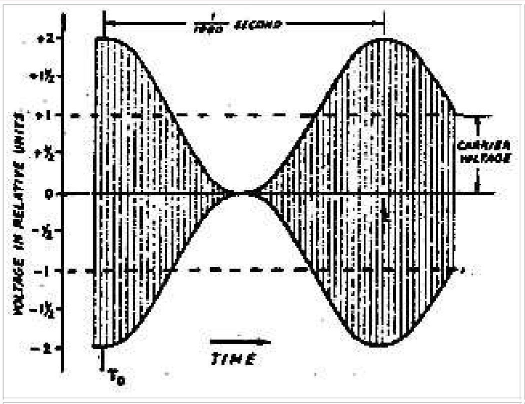


Fig 1. Envelope of carrier 100% modulated by a 1000 Hz sine wave.

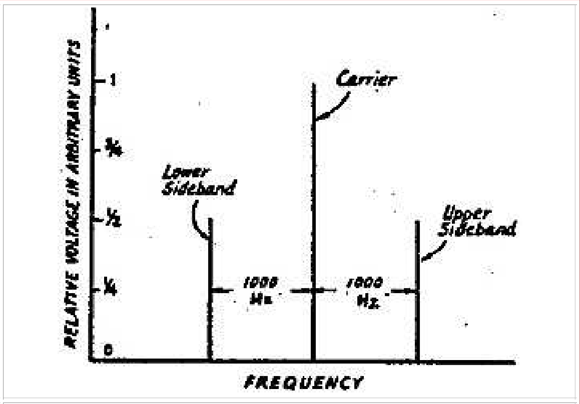


Fig 2 - Example of 100% modulation of a carrier by a single tone of 1000 Hz

Since power is proportional to the square of voltage, the proportion of power in each side frequency is equal to 0.5 squared or a quarter of that in the carrier. If the carrier is modulated 100 per cent by a complex waveform of many frequencies, two sidebands of frequencies are created, each with power equal to a quarter of the carrier power. The significance of all this is shown by considering a carrier of 100 watts modulated 100 per cent. Additional power of 25 watts in each sideband is also transmitted, making a total power of 150 watts. The intelligence transmitted in the complete modulated signal is contained in the sidebands and only one of these is needed to support this intelligence. Here we see a reason why single sideband (SSB) transmission is used in preference to transmitting the basic AM. (25 watts of SSB is just as effective as 150 watts of AM carrier plus sidebands).

A further consideration is the bandwidth taken up by the amplitude modulated signal . 'To transmit good quality speech, audio frequencies in the range of around 200Hz to 2500Hz must modulate the carrier. Hence the sidebands extend from 2500Hz below the carrier frequency to 2500Hz above, requiring a complete bandwidth of 5000Hz (refer figure 3). If one sideband and the carrier are suppressed, as in the SSB system, bandwidth is reduced to 2300Hz, less than half that of the AM signal. This means that the receiver bandwidth can be halved and more signals can be fitted in a given bandspace to be received without interference. Suppression of continuous transmitted carrier on adjacent signals also results in improved reception as heterodyne whistles are eliminated. These whistles are often a problem on a crowded band of AM signals.

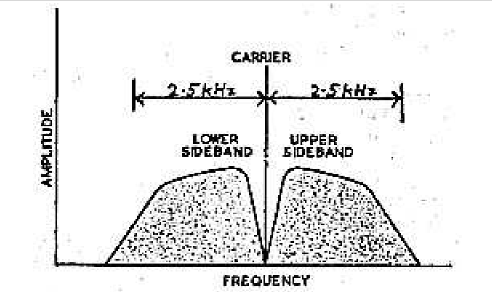


Fig 3 - Amplitude/lfrequency relationships of carrier and sidebands with 100% speech modulation.

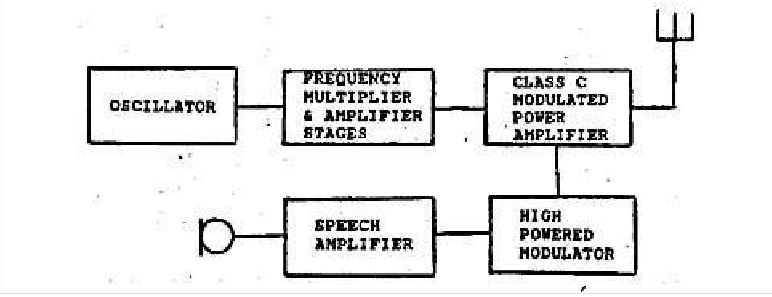


Fig 4 - High level (plate modulated) amplitude modulation transmitter.

The AM Transmitter

In replacing the AM system with the SSB system, some significant savings are achieved in the size of the RF power amplifier and in the power supply which feeds the amplifier. Let us consider an AM transmitter which is to deliver 100 watts of carrier power. High level or plate modulation (as shown in figure 4) is to be used and, in this system, our final RF power amplifier can run at the highest possible efficiency in Class C operation. In this form of operation, the output tuned circuit (called a tank circuit) is pumped to maintain it in oscillation by pulses from the RF power amplifier. The amplifier is biased to allow only a small portion of the RF drive sine waveform to be amplified, so providing the pumping pulse. In this form of operation (Class C), the amplifier can be made to operate much more efficiently than as a linear amplifier which must reproduce the complete sine wave fed into it. Efficiencies in the order of 70 to 80 per cent can be expected from a Class C amplifier, and for our transmitter we will assume a value of 75 per cent. Our total input power to the transmitter is therefore 133 watts of which 100 watts is radiated and 33 watts is dissipated in the amplifier in heat. To supply the amplifier, we need a power supply which can deliver a power of 133 watts and we need an amplifier valve (or perhaps a transistor) which can dissipate a continuous power of 33 watts.

To 100 per cent modulate our transmitter, we also require 50 per cent of extra power for the sidebands, and this is 50 per cent of the input power 133 watts, not the 100 watts of output power. The extra power is supplied as 67 watts of audio from the modulator output stage. The modulator stage runs as a high efficiency power amplifier in Class AB or Class B and a practical efficiency might be as high as 60 per cent which we will assume. At 100 per cent modulation continuous tone, DC load on the modulator power supply is 112 watts and the modulator valves or transistors must dissipate 45 watts. The output power on speech is only about 20 per cent of the peak power which gives 100 per cent modulation. However, the average DC power is somewhat higher than this because of the zero signal standing current into the amplifier. In Class AB, the power input at zero signal might be in the order of one quarter to one third of the input power at peak output, and for our case, around 28 to 37 watts. All in all, the modulator amplifier must dissipate an average power of around 30 to 40 watts, and the modulator power supply must be able to supply a varying load which swings between say 30 to 112 watts.

To summarise this, our 100W AM transmitter requires the following: :

- RF power amplifier - 33 watts continuous dissipation
- Modulators - Average dissipation 30 to 40 watts
- RF amp power supply - 133 watts continuous
- Modulator power supply - Swinging load 30 to 112 watts

As an alternative, we could use low level modulation to eliminate the high power modulator and its power supply. Either grid modulation is used or an RF driver stage is modulated and followed up with a linear RF final power amplifier as shown in figure 5. In either case, the final amplifier, being a linear stage, must be operated at a much lower efficiency than in class C. Hence, higher dissipation amplifiers are needed and a larger power supply. What is gained in reduced modulator power is lost in extra dissipation power in the final RF amplifier.

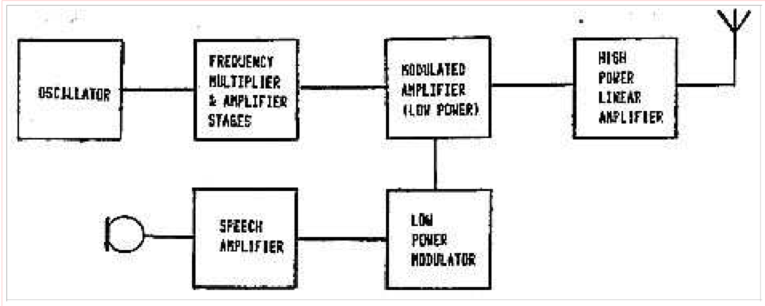

OSCILLATOR
FREQUENCY MULTIPLIER & AMPLIFIER STAGES
MODULATED AMPLIFIER (LOW POWER)
HIGH POWER LINEAR AMPLIFIER
SPEECH AMPLIFIER
LOW POWER MODULATOR

Fig 5 Low Level Amplitude Modulation Transmitter.

The SSB Transmitter

A typical single sideband transmitter is shown in figure 6. The SSB signal is generated in the low-level stages of the transmitter. A balanced modulator is used to balance out the fixed carrier of 9 MHz, leaving a double sideband suppressed carrier signal. This is fed through a fixed frequency narrow band filter designed to slice off one of the two sidebands. The remaining sideband is mixed with a variable frequency oscillator (VFO) to produce the SSB signal at the required operational frequency. All stages following the modulator (including the final RF power amplifier) must be operated in a linear mode. Since the final amplifier is linear, it cannot be operated in Class C, and its power efficiency is lower than that obtainable in a high level modulated AM transmitter. Before considering this to be a disadvantage, we must first examine actual powers involved.

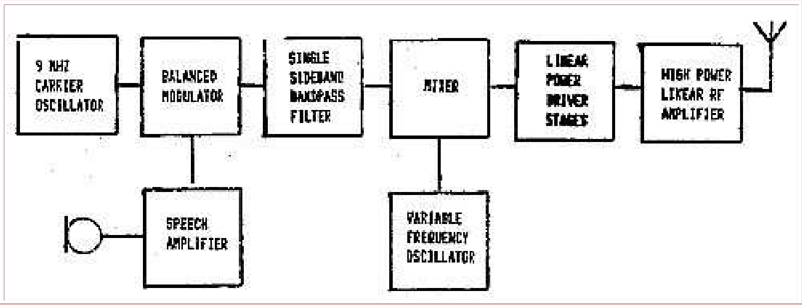


Fig 6 - Typical single sideband transmitter.

It was pointed out earlier that 25 watts of single sideband was equally effective as a 100W carrier AM signal fully modulated. To make comparison with our AM transmitter, we use 25 watts of SSB. Our final linear RF amplifier (in say Class AB) might he expected to have a typical efficiency of 50 per cent. At this percentage, input power is 50 watts and hence our power supply must deliver 50 watts and our amplifier must dissipate 25 watts. This is a large improvement on the 133 watts of input power and 33 watts of dissipation quoted for our AM transmitter, but our gain is even better than this. The 50 watts input to the SSB amplifier on speech is our peak envelope power (PEP). As we discussed earlier, the average power into a Class AB amplifier on speech is much less than this and possibly in the region of 30 percent of the peak value. Taking this percentage, the average input power is only 15 watts, with the average dissipation perhaps half of the 15 watts. (The average dissipation will depend much on what standing current is run in the no-signal condition between speech syllables).

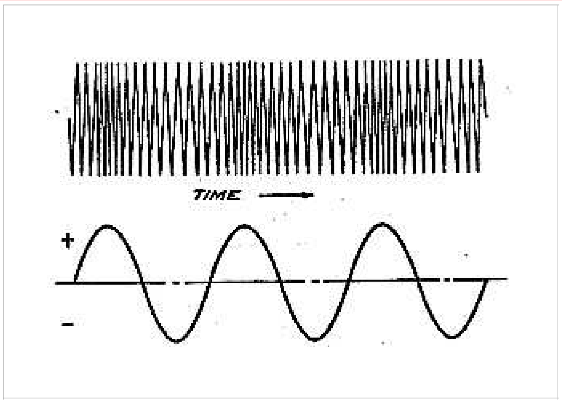
Now to summarise the SSB transmitter:

- RF power amplifier - 7.5 watts average dissipation
- RF amp power supply - Average power load 15 watts with regulation to allow for short duration peaks of 50 watts.
- High power modulator not required.

Comparing this to the AM transmitter, previously described, we see that SSB offers a considerable reduction in the ratings and size of components used in the final stages of the transmitter. Even though the SSB circuitry is a little more complicated, the SSB transmitter can be made more compact than the AM unit of equal effective power.

Frequency Modulation

In frequency modulation (FM), the frequency of the carrier wave is varied as a function of the instantaneous voltage of the modulating signal. This is illustrated in figure 7. The amount of frequeney shift off the centre frequency is called the frequency deviation. A peak deviation of 5kHz (such as used in amateur radio systems) means that the carrier frequency is shifted in one direction a maximum of 5kHz by the positive going peaks of the modulating signal and shifted in the opposite direction a maximum of 5kHz by the negative going peaks of the modulating signal. Total frequency swing is thus 10kHz.



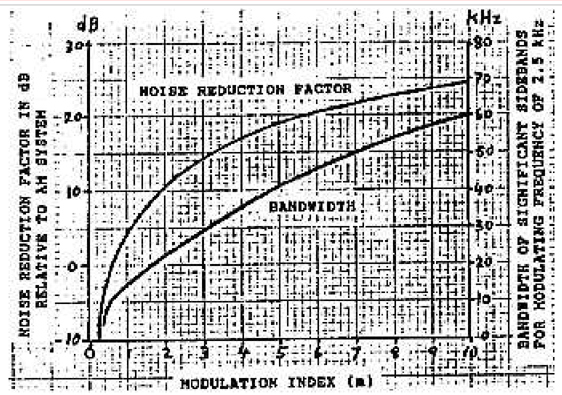


Fig 7 - A frequency-modulated signal in which the frequency of the signal varies in accordance with the level of the modulating audio voltage. At points where the audio voltage is positive, the frequency is high, whilst at points where the audio voltage is negative, the frequency is low	**Fig 8 - Comparison of bandwidth and noise reduction factor in FM system for different values of modulation index at a modulating frequency of 2.5kHz**

Modulation index is defined as the ratio of frequency deviation to modulating frequency producing the deviation. If a 1kHz modulating signal produces 5kHz of deviation. the modulation index is equal to 5. Considering a maximum speech frequency of 2.5kHz, the modulation index equals 2 if the carrier frequency is driven to a maximum deviation of 5kHz by that particular speech frequency component.

The FM receiver is designed to be insensitive to amplitude variation in the RF signal it receives. As random incoming noise is received essentially as a voltage of fluctuating amplitude, the receiver on FM has a signal to noise ratio advantage over an AM receiver, given received signals of equal carrier amplitude. The degree of that advantage is dependent on the modulation index which is used and this is illustrated in figure 8 showing noise reduction factor in dB as a function of the index. The diagram shows that to gain advantage, the modulation index must be greater than 0.6, and the higher the value of the index, the greater is the noise reduction factor. In comparing the FM and AM systems, equal receiver audio bandwidth is assumed.

All this is fine except that the FM signal has sidebands much more complicated than the AM signal, and which theoretically extend infinitely either side of the carrier frequency. In practice, we need only to consider the sideband frequencies which are of significant level. The bandwidth of the significant sidebands increases both as the modulation index is increased and as the modulating frequency is increased. The second curve in figure 8 plots the bandwidth of the significant sidebands as a function of modulation index for a modulating frequency of 2.5kHz, chosen as the maximum speech frequency. Using both curves, we see that to get a 10dB signal to noise ratio advantage we need a modulation index equal to 2. However, to achieve this, we take up a bandwidth of around 22kHz.

So here is the answer to why FM is restricted essentially to the VHF and UHF bands. FM gives us a signal to noise ratio advantage over AM, but it takes up more bandwidth and much more than we are able to accommodate in the restricted bandspace of our HF bands. More bandspace is available on the VHF and UHF bands, allowing us to use FM as a popular mode of transmission. On two metres, for example, we use 25kHz channel spacing to accommodate the wide bandwidth FM signals.

Frequency modulation is actually allowed on the HF bands, but bandwidth is restricted to 6kHz, limiting the modulation index to around 0.4 at 2.5kHz modulating frequency. With this restriction, performance cannot be expected to be any better than AM, and not as good as SSB.

The FM Transmitter

As in the SSB transmitter, modulation of the FM signal is carried out in the low level stages of the transmitter (refer figure 9). However, unlike the SSB transmitter, linear amplification is not required following modulation and the following stages (including the final amplifier) can be run at highest efficiency in Class C operation. The final amplifier efficiency is thus similar to that of an AM transmitter using high level modulation. Of course, a high power modulator is not required as in the AM transmitter so that the high power circuits in the FM transmitter are less complicated than for the AM transmitter of similar carrier power.

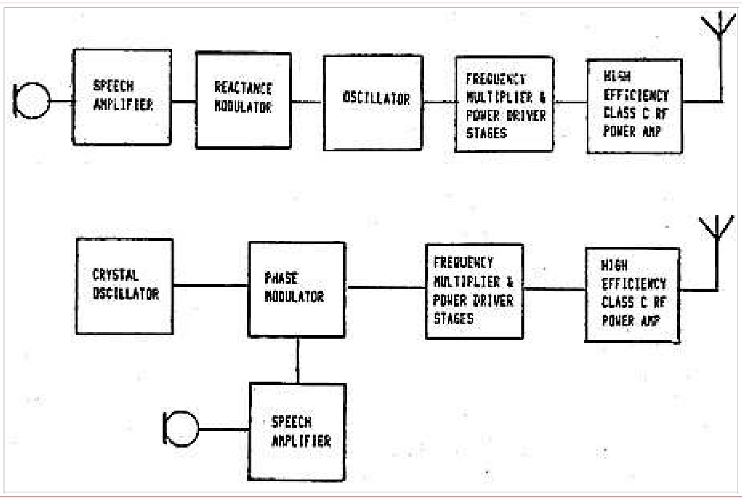


Fig 9 - Typical frequency modulation transmitters.

Squelch

Mobile radio systems essentially use the VHF and UHF spectrums where bandspaoe is less of a premium than in the HF spectrum and where the VHF and UHF frequencies are more suited to the short range communication required. Most of the mobile radio networks now use FM to gain the signal to noise ratio advantage over AM (which was used in earlier systems). SSB has not generally been used and we offer one very good reason for this. A desirable requirement in a mobile vehicle is to maintain the radio silent when no signal is being received. To do this, a "squelch" circuit is used which turns on the receiver audio stages only when a carrier is being received. Of course, a SSB signal has no carrier and the sidebands are only sent on speech syllables. Operation of squelch on these could be erratic, particularly in the presence of noise. So the point is made that SSB is at a disadvantage for systems, such as mobile radio, where a simple squelch system is desired.

CW

Of all the modes of transmission used in amateur radio, what we call continuous wave (CW) transmission is the simplest to generate. The RF continuous wave is simply turned on and off by some manual or automatic keying device to transmit an intelligible code. If we examine the transmitted waveforms with a spectrum analyser, in what we have previously referred to as the frequency domain, we see that the signal is another example of amplitude modulation. The display shows a carrier frequency with sidebands formed from the modulating signal which is a modified square (or rectangular) waveform with its fundamental frequency (or frequencies) set by the keying speed.

Of course, the squared keying waveform must be band limited as a perfect square wave has infinite odd harmonics which would produce sidebands of infinite width. Excessive sidebands are heard as key clicks at frequencies extended either side of the operating frequency. To prevent this, the keying circuit is fed through a low pass filter to limit the harmonics and hence the bandwidth radiated. For satisfactory aural reception of Morse code, references recommend that at least the third harmonic should be transmitted for non-fading conditions and both third and fifth harmonics for fading conditions.

The rate per second at which the keying signal changes its state, either space to mark or mark to space, is called the baud rate. For Morse code, the baud rate has been quoted as approximately equal to the Morse speed in words per minute (WPM) divided by 1.2. For 20wpm speed, the baud rate is thus 16.7 and the fundamental frequency is half that, or 8.3Hz.

Assuming we radiate up to the fifth harmonic of the keying frequency, each sideband is 5 x 8.3 = 41.7Hz wide, and the total bandwidth is twice this, or 83.4Hz. Because of the narrow bandwidth, many more Morse code CW signals can be fitted in a given bandspace than any form of voice modulation. To take full advantage of this fact, a good receiver for CW should have a crystal filter, or some other means, to restrict its bandwidth to hundreds of Hertz.

The CW transmitter is the simplest of all. The RF circuits can all run at maximum efficiency in Class C, and no high power transistor or valve modulator is required, as the modulator is the simple keying circuit. The final amplifier stage in a CW transmitter can usually be run at higher power than the high level modulated final amplifier in an AM speech transmitter. In the latter, instantaneous voltage applied to the modulated amplifier is doubled at peaks of modulation and the stage must be rated to withstand this. Furthermore, the power dissipation is continuous whereas, in the CW transmitter, average power is reduced by the on/off keying operation.

Frequency Shift Keying

In automatic radio telegraphy and digital data transmission systems, we normally use frequency shift keying, as this is another form of frequency modulation which gives improved signal to noise ratio. Radio teletype (RTTY) as used in amateur radio generally runs at a baud rate of 45.45 (and sometimes 50) with a Frequency shift of 170Hz. The fundamental modulating frequency is thus 45.45/2 = 22.73Hz and the frequency deviation is 170/2 = 85.5Hz. The modulation index at the fundamental frequency is therefore 85.5/22.73 = 3.76 which, from figure 6, gives a signal to noise ratio improvement of 16dB over AM or amplitude keying.

From references, the minimum bandwidth for the teletype is given as being equal to baud rate plus frequency shift multiplied by 1.2. Using this, minimum bandwidth = 45.45 + 170 x 1.2 = 249.45 (let's say 250).

In amateur radio teletype we use audio frequency shift keying (AFSK) and use standard frequencies of 2125Hz for mark and 2295 for space or 1275Hz for mark and 1445 for space. By feeding these tones into the audio input circuit of our SSB transmitter, the RF single sideband generated appears as if we were simply shifting a carrier at 170Hz of shift.

Using FM equipment at VHF, the tones are again fed to the audio input, but in this case we have a frequency modulated audio sub-carrier in turn frequency modulating the RF carrier. In this case, derivation of the significant bandwidth is a little more complicated. The audio bandwidth is calculated as before as equal to 250Hz. The highest frequency in the audio tones is then determined and, for the 2125/2295Hz tones, this is worked out by taking the average of these frequencies and adding it to half the audio bandwidth. The result is 2334.5Hz. The radiated FM bandwidth is now worked out by adding the highest audio frequency result to the frequency deviation used and multiplying by 2. For 5kHz deviation, the significant bandwidth works out to 14.669kHz.

Using frequency shift of the carrier, as is achieved by feeding the audio tones into the SSB transmitter, the RTTY transmission is a narrow-band mode similar in bandwidth to Morse with CW transmission. Feeding the tones into an FM transmitter (or, for that matter, a double sideband with carrier AM transmitter), the RTTY is a wide-band mode similar to speech.

Packet radio systems operate at much higher baud rates than RTTY, and baud rates vary from 300 (often used on the HF bands) to 9600 (for meteor-scatter and satellite communications where access time is limited). A 300-baud system using an RF carrier shift of 200 Hz has a modulation index of 0.67 and requires a bandwidth of 540Hz. Higher baud rates (typically 1200 on the VHF band) call for bandwidths comparable with speech or greater.

Television

To reproduce, with high definition, the picture elements in a PAL system. TV picture, video components approaching 5MHz must modulate an RF carrier. Together with the sound, which also modulates a carrier, a total bandwidth of around 7MHz is needed. In a standard PAL system, the video signal amplitude modulates the vision carrier, and one complete sideband is fully transmitted. For the other sideband, bandspace is reduced by transmitting the low frequency modulation components of the sideband up to around 1.25MHz. The system is called vestigial sideband transmission. Whilst only one complete sideband is needed to convey the signal information, there is a problem in making a sideband filter without phase shift around its cut-off frequency. Reproduction of the TV picture is seriously affected by phase shift in the low frequency components and, hence, the sideband cut-off is shifted up to well above the frequencies which are most affected.

To take advantage of the improved signal to noise ratio of FM and the high quality sound reproduction which can be achieved, frequency modulation is used in the TV broadcasting system. The system also minimises interference from the amplitude modulated video signals. The sound carrier is spaced exactly 5.5MHz from the vision carrier and, in the TV receiver, the two carriers beat together to form the 5.5MHz sound IF channel.

The standard PAL signal format is shown in figure 10. This shows the vestigial sideband arrangement and the difference frequency of 5.5MHz between the vision and sound carriers. For colour TV, an additional 4.43MHz colour subcarrier is superimposed on the video signal and this appears as a single sideband component 4.43MHz above the vision carrier. As seen by the diagram, the system takes up a bandwidth of 7 MHz, some hundreds of times the bandwidth used by even our widest speech modes. Hence, amateur TV is confined to the UHF bands where more bandspace is available. Amateur TV may also operate as a double sideband system and, as seen in figure 11. This system takes up a bandspace of over 11MHz. Fast-scan TV (as we call the system described) is clearly not suitable for the lower frequency bands. However, there is a group of enthusiasts who transmit slow-scan TV which requires a more modest bandspace.

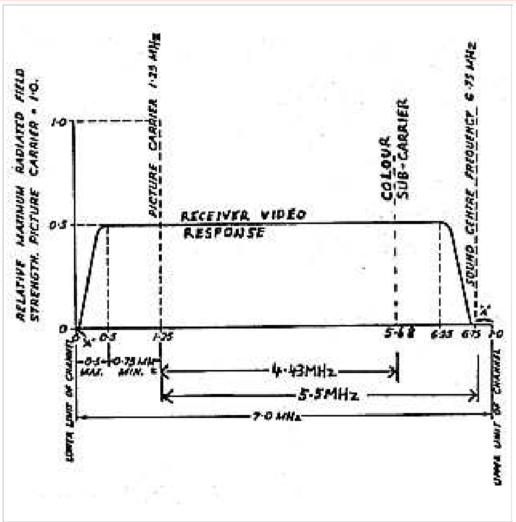


Fig 10 - PAL television signal format.

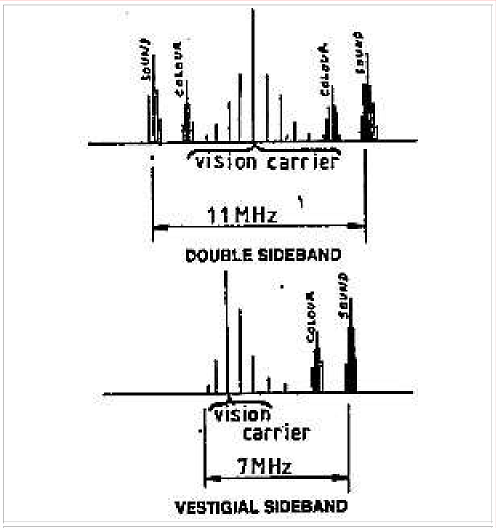


Fig 11 - Comparison of double sideband and vestigial sideband TV signal formats

In the PAL TV system, the complete picture or frame is scanned 25 times per second, and each frame is made up of 625 lines (including those which are not seen during the scan retrace). In slow-scan TV, the scanning rate is slowed down to one frame in a number of seconds, and the number of lines in a frame is reduced. Images appear as a series of still pictures which change with the movement of the televised object rather than as a continuous moving picture. Bandwidth is reduced both by the reduction in scanning speed and by transmitting picture elements of lower definition. Using slowscan TV, signal bandwidths comparable with speech can he achieved, and hence it is feasible to transmit on the HF bands.

Summary

Various forms of modulation and how they are applied to the modes of transmission have been discussed. In amplitude modulation, there is much to be gained in suppressing one of the sidebands and the carrier signal. Hence single sideband transmission has been phased in over the years in preference to the basic AM system. Advantages are a reduction in bandspace, elimination of heterodyne whistles heard on the band and more effective use of RF power generated in the transmitter. For a given effective signal, a lower power rating in the RF amplifier and a smaller power supply are achieved.

Frequency modulation has a signal to noise ratio advantage over amplitude modulation, but to gain the advantage, the modulation index must be greater than 0.6 with a resultant bandwidth on speech considerably greater than that required for amplitude modulation. Because of this, FM is essentially used on the VHF/UHF bands where the wider bandwidth can be better accommodated.

CW transmission, as we know it, is a form of amplitude modulation, and bandwidth must be allowed for the sidebands generated from the keying characters. At hand-keying speeds, quite low fundamental frequency components are generated. Hence CW signals occupy a narrow bandspace provided the keying signal is adequately filtered to remove higher order harmonics.

Frequency shift keying, as used in RTTY and digital data systems such as packet radio, is a form of frequency modulation in which significant FM sidebands are generated. The bandwidth of these sidebands is determined by the baud rate (or modulating frequency component) and the amount of frequency shift (or frequency deviation).

Fast-scan television has such a wide bandwidth that amateur experimentation is restricted to the UHF bands. Slowscan television, as used in amateur radio, has a bandwidth comparable with speech and can be used on the HF bands.

For further reading, some excellent material on modulation and data transmission systems can be found in the latest issues of the ARRL Handbook. References to this and other sources of information are included following.

References

1. ARRL Handbook 1989. Chapters on "Modulation & Demodulation", "Digital Communications", "Voice Communications".

2. R F Dannecker VK4ZFD "The FM System", Amateur Radio, Dec 1969.

3. A H Kaye. "Frequency Modulation, Principles & Equipment Fundamentals", Amateur Radio, Nov 1947.

4. John Kennewell & Kenneth Brown. "Frequency Modulation & Sidebands", Electronics Australia, Nov 1983.

5. Bryan Maher. "FM Bandwidth & Frequency Response", Electronics Australia, Jan 1990

6. Lloyd Butler VK5BR. "Spectrum Analyser Waveforms" Amateur Radio, Sept 1987 page 24.

7. Australian Amateur Call Books 1983/84 or 1984/85. WIA publication, "Amateur Fast-Scan Television'.

Back to HomePage