FM Demodulator , Ratio and Foster-Seeley FM detectors

There are a number of circuits that can be used to demodulate FM. Each type has its own advantages and disadvantages, some being used when receivers used discrete components, and others now that ICs are widely used.

Slope detection

The very simplest form of FM demodulation is known as slope detection or demodulation. It simply uses a tuned circuit that is tuned to a frequency slightly offset from the carrier of the signal. As the frequency of the signal varies up and down in frequency according to its modulation, so the signal moves up and down the slope of the tuned circuit. This causes the amplitude of the signal to vary in line with the frequency variations. In fact at this point the signal has both frequency and amplitude variations. The final stage in the process is to demodulate the amplitude modulation and this can be achieved using a simple diode circuit. One of the most obvious disadvantages of this simple approach is the fact that both amplitude and frequency variations in the incoming signal appear at the output. However the amplitude variations can be removed by placing a limiter before the detector. Additionally the circuit is not particularly efficient as it operates down the slope of the tuned circuit. It is also unlikely to be particularly linear, especially if it is operated close to the resonant point to minimise the signal loss.

Ratio and Foster-Seeley FM detectors

When circuits employing discrete components were more widely sued, the Ratio and Foster-Seeley detectors were widely used. Of these the ratio detector was the most popular as it offers a better level of amplitude modulation rejection of amplitude modulation. This enables it to provide a greater level of noise immunity as most noise is amplitude noise, and it also enables the circuit to operate satisfactorily with lower levels of limiting in the preceding IF stages of the receiver.

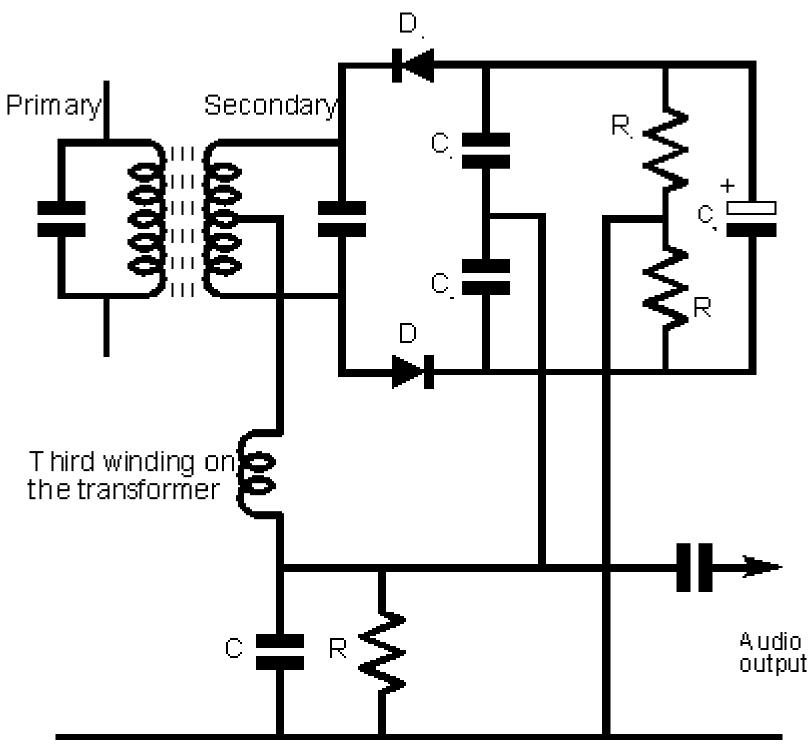
The operation of the ratio detector centres around a frequency sensitive phase shift network with a transformer and the diodes that are effectively in series with one another. When a steady carrier is applied to the circuit the diodes act to produce a steady voltage across the resistors R1 and R2, and the capacitor C3 charges up as a result.

The transformer enables the circuit to detect changes in the frequency of the incoming signal. It has three windings. The primary and secondary act in the normal way to produce a signal at the output. The third winding is un-tuned and the coupling between the primary and the third winding is very tight, and this means that the phasing between signals in these two windings is the same.

The primary and secondary windings are tuned and lightly coupled. This means that there is a phase difference of 90 degrees between the signals in these windings at the centre frequency. If the signal moves away from the centre frequency the phase difference will change. In turn the phase difference between the secondary and third windings also varies.

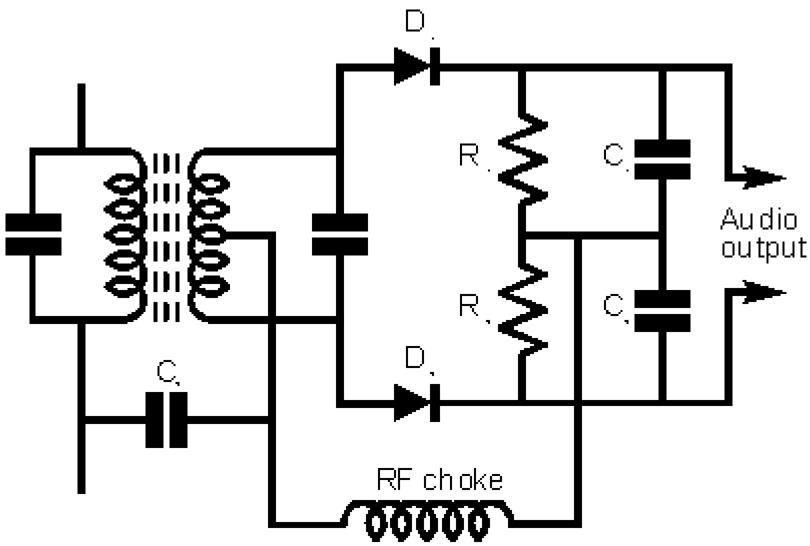
When this occurs the voltage will subtract from one side of the secondary and add to the other causing an imbalance across the resistors R1 and R2. As a result this causes a current to flow in the third winding and the modulation to appear at the output.

The capacitors C1 and C2 filter any remaining RF signal which may appear across the resistors. The capacitor C4 and R3 also act as filters ensuring no RF reaches the audio section of the receiver.



The ratio detector

The Foster Seeley detector has many similarities to the ratio detector. The circuit topology looks very similar, having a transformer and a pair of diodes, but there is no third winding and instead a choke is used.



The Foster-Seeley detector

Like the ratio detector, the Foster-Seeley circuit operates using a phase difference between signals. To obtain the different phased signals a connection is made to the primary side of the transformer using a capacitor, and this is taken to the centre tap of the transformer. This gives a signal that is 90 degrees out of phase.

When an un-modulated carrier is applied at the centre frequency, both diodes conduct, to produce equal and opposite voltages across their respective load resistors. These voltages cancel each one another out at the output so that no voltage is present. As the carrier moves off to one side of the centre frequency the balance condition is destroyed, and one diode conducts more than the other. This results in the voltage across one of the resistors being larger than the other, and a resulting voltage at the output corresponding to the modulation on the incoming signal.

The choke is required in the circuit to ensure that no RF signals appear at the output. The capacitors C1 and C2 provide a similar filtering function.

Both the ratio and Foster-Seeley detectors are expensive to manufacture. Wound components like coils are not easy to produce to the required specification and therefore they are comparatively costly. Accordingly these circuits are rarely used in modern equipment.

Quadrature FM detector

Another form of FM detector or demodulator that can be these days is called the quadrature detector. It lends itself to use with integrated circuits and as a result it is in widespread use. It has the advantage over the ratio and Foster-Seeley detectors that it only requires a simple tuned circuit.

For the quadrature detector, the signal is split into two components. One passes through a network that provides a basic 90 degree phase shift, plus an element of phase shift dependent upon the deviation and into one port of a mixer. The other is passed straight into another port of the mixer. The output from the mixer is proportional to the phase difference between the two signals, i.e. it acts as a phase detector and produces a voltage output that is proportional to the phase difference and hence to the level of deviation on the signal.

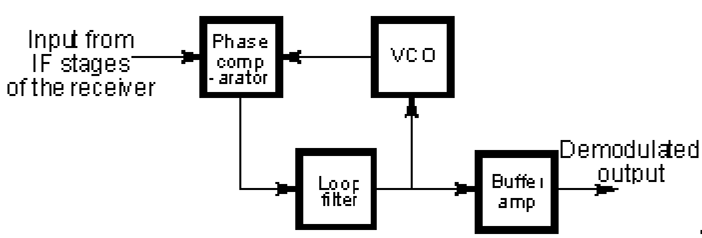
The detector is able to operate with relatively low input levels, typically down to levels of around 100 microvolts and it is very easy to set up requiring only the phase shift network to be tuned to the centre frequency of the expected signal. It also provides good linearity enabling very low levels of distortion to be achieved.

Often the analogue multiplier is replaced by a logic AND gate. The input signal is hard limited to produce a variable frequency pulse waveform. The operation of the circuit is fundamentally the same, but it is known as a coincidence detector. Also the output of the AND gate has an integrator to "average" the output waveform to provide the required audio output, otherwise it would consist of a series of square wave pulses.

Phase locked loop (PLL)

Another popular form of FM demodulator comes in the form of a phase locked loop. Like the quadrature detector, phase locked loops do not need to use a coil, and therefore they make a very cost effective form of demodulator.

The way in which they operate is very simple. The loop consists of a phase detector into which the incoming signal is passed, along with the output from the voltage controlled oscillator (VCO) contained within the phase locked loop. The output from the phase detector is passed into a loop filter and then sued as the control voltage for the VCO.



Phase locked loop (PLL) FM demodulator

With no modulation applied and the carrier in the centre position of the pass-band the voltage on the tune line to the VCO is set to the mid position. However if the carrier deviates in frequency, the loop will try to keep the loop in lock. For this to happen the VCO frequency must follow the incoming signal, and for this to occur the tune line voltage must vary. Monitoring the tune line shows that the variations in voltage correspond to the modulation applied to the signal. By amplifying the variations in voltage on the tune line it is possible to generate the demodulated signal.

It is found that the linearity of this type of detector is governed by the voltage to frequency characteristic of the VCO. As it normally only swings over a small portion of its bandwidth, and the characteristic can be made relatively linear, the distortion levels from phase locked loop demodulators are normally very low.

Frequency Vs Phase Modulation:

The difference between FM & PM in a digital oscillator is that FM is added to the frequency before the phase integration, while PM is added to the phase after the phase integration. Phase integration is when the old phase for the oscillator is added to the current frequency (in radians per sample) to get the new phase for the oscillator. The equivalent PM modulator to obtain the same waveform as FM is the integral of the FM modulator. Since the integral of sine waves are inverted cosine waves this is no problem. In modulators with multiple partials, the equivalent PM modulator will have different relative partial amplitudes. For example, the integral of a square wave is a triangle wave; they have the same harmonic content, but the relative partial amplitudes are different. These differences make no difference since we are not trying to exactly recreate FM, but real (or nonreal) instruments.

The reason PM is better is because in PM and FM there can be non-zero energy produced at 0 Hz, which in FM will produce a shift in pitch if the FM wave is used again as a modulator, however in PM the DC component will only produce a phase shift. Another reason PM is better is that the modulation index (which determines the number of sidebands produced and which in normal FM is calculated as the modulator amplitude divided by frequency of modulator) is not dependant on the frequency of the modulator, it is always equal to the amplitude of the modulator in radians. The benefit of solving the DC frequency shift problem, is that cascaded carrier-modulator pairs and feedback modulation are possible. The simpler calculation of modulation index makes it easier to have voices keep the same harmonic structure throughout all pitches.