

Demodulators and detectors

In communications equipment, “detection” is synonymous with demodulation, the process of recovering information from the received signal. The term detector is also used for circuits designed to measure power, such as square-law microwave power detectors. In this chapter we discuss various AM, FM, and power detector circuits. The demodulator is the last module in the cascade of circuits that form a receiver and at this stage the frequencies (IF or baseband) are relatively low. For this reason the detector (and, sometimes, the final IF bandpass filter) was the first receiver section to evolve from analog to digital processing, notably in broadcast television receivers. Receivers for ordinary AM and FM have, for the most part, continued to use traditional analog detectors, but receivers for the newer digital radio formats can also use their digital signal processors to demodulate traditional AM and FM broadcasts. Demodulators for OFDM and CDMA digital modulation formats are discussed in Chapter 22.

18.1 AM detectors

There are two basic types of AM detector. An envelope detector uses rectification to produce a voltage proportional to the amplitude of the IF voltage. A “product” detector multiplies the IF signal by a reconstituted version of the carrier. This detector is a mixer (see Chapter 5), producing sum and difference frequencies. The sum component is filtered away. The difference frequency component, at $f = 0$ (baseband), is proportional to the amplitude of the IF voltage.

18.1.1 AM diode detector

The classic diode envelope detector circuit for AM is shown in Figure 18.1. The input signal voltage, $V_{\text{sig}}\cos(\omega t)$, is usually provided by a tuned transformer at the output of the IF amplifier. This tuned circuit forms part of the IF bandpass filter.

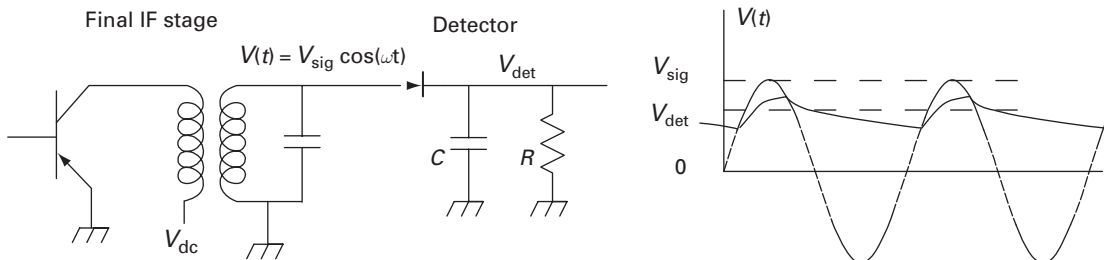


Figure 18.1. Diode envelope detector.

The diode and the parallel RC form a fading-memory peak detector. Except for the resistor, R , the output capacitor would remain charged to the maximum peak voltage of the input sine wave. The resistor provides a discharge path so that the detector output, V_{det} , can follow a changing peak voltage (AM). Since the input sine wave, an RF signal, has a much higher frequency than the amplitude modulation frequencies, the RC time constant can be made large enough so that the droop between charge pulses is much less than indicated in Figure 18.1. (Of course the time constant must also be small enough that output voltage can accurately follow a rapidly changing modulation envelope.) This detector, or any other envelope detector, is known as a *linear detector* since its output voltage is linearly proportional to the amplitude of the input sine wave.

Analysis assuming an ideal rectifier

Note that the detector of Figure 18.1 is identical to a simple half-wave rectifier capacitor-input power supply. As with the power supply, this circuit has poor regulation with respect to a changing load. But here we have a constant load resistance R (which we will assume includes the parallel input resistance of the subsequent audio amplifier). The equivalent circuit is shown in Figure 18.2. If the diode is modeled as a perfect rectifier (zero forward resistance and infinite reverse resistance) the analysis of this circuit is straightforward. The value of the

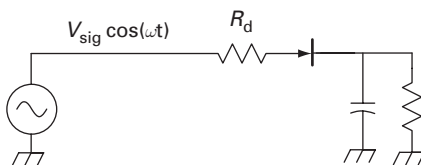


Figure 18.2. Diode envelope detector equivalent circuit.

diode's forward resistance, R_d , can be increased to account for source resistance. But here the high- Q resonant circuit at the detector input forces the waveform to remain sinusoidal. If we assume that C is large enough to make the output ripple negligible compared to the output voltage, the output voltage, V_{det} , can be calculated by noting that V_{det}/R must equal the average current through the diode, which is the average of $(V - V_{det})/R_d$ during the part of the input cycle

when this expression is positive. The result is that V_{det} is proportional to the source voltage. (Curves showing output voltage vs. ωCR for various values of R_d/R are found in the power supply chapters of many handbooks.) The ratio of V_{det} to the peak source voltage is known as the detector efficiency. For a typical AM detector, $R \geq 10R_d$ and $\omega CR \geq 100$. This gives a detector efficiency greater than 65% and an rms ripple less than about 1% of the dc output. With the assumed ideal rectifier, however, R could be any value. The analysis in the following section shows the limitations imposed on R by a real diode.

Analysis with a real diode

Here we will discard the perfect rectifier in favor of the standard diode for which $I_{\text{diode}} = I_s \exp(V_{\text{diode}}/V_T - 1)$, where I_s is the reverse saturation current and V_T is the so-called thermal voltage, 0.026 volts.¹ In the equivalent circuit of Figure 18.2, we will now let R_d be zero, i.e., we will assume that any voltage drop across the diode's bulk resistance is negligible compared to the drop across the junction. As before, the analysis to find V_{det} consists in equating V_{det}/R to the average current through the diode:

$$\frac{V_{\text{det}}}{R} = \langle I_{\text{diode}} \rangle = \frac{1}{2\pi} \int_0^{2\pi} I(\theta) d\theta \quad (18.1)$$

where

$$I(\theta) = I_s \{ \exp[(V \cos \theta - V_{\text{dc}})/V_T] - 1 \}. \quad (18.2)$$

This pair of equations is equivalent to

$$\frac{V_{\text{det}}}{0.026} = \ln \left(\frac{V_s I(V)}{2\pi(V_{\text{det}} + V_s)} \right), \quad (18.3)$$

where V_s , a “saturation voltage,” is defined by $V_s = I_s R$ and

$$I(V) = \int_0^{2\pi} \exp(V \cos \theta / 0.026) d\theta. \quad (18.4)$$

Using a desk-top computer math utility to solve Equation (18.3) results in a set of curves (Figure 18.3) showing V_{det} vs. V for various values of V_s . Note that the detector output is very nonlinear for low-amplitude input signals when V_s is less than about 0.01 V. Suppose, then, that we pick $V_{\text{sat}} = 0.01$ V. A germanium diode or zero-bias Schottky diode might have a saturation current of 10^{-6} A, which would

¹ The thermal voltage is given by kT/e , where e is the charge of an electron, k is Boltzmann's constant, and T is an assumed temperature, 300 K.

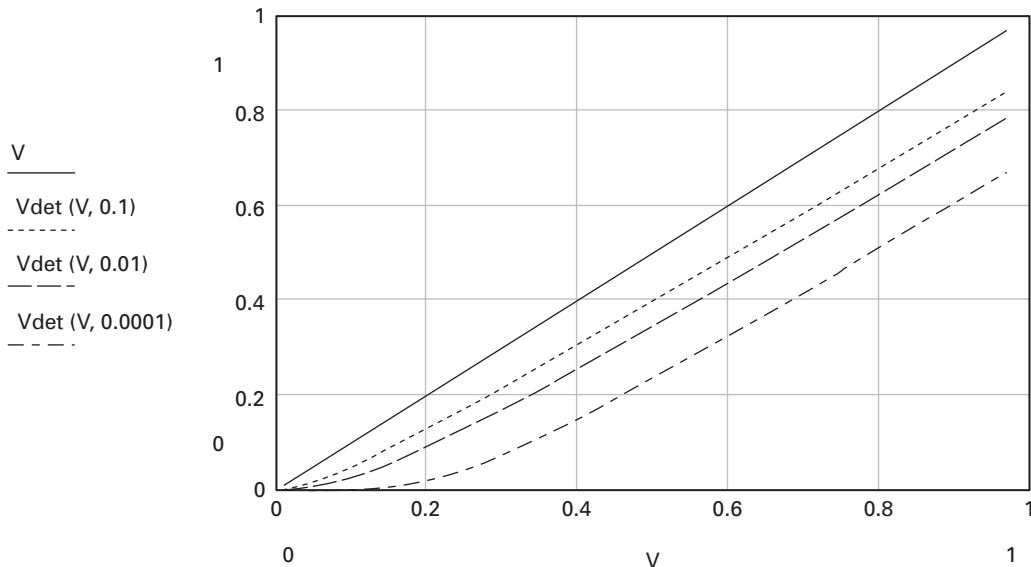


Figure 18.3. Detector output vs. input voltage for several values of $V_s = I_s R$.

then require that $R = 0.01/10^{-6} = 10 \text{ k ohms}$, a convenient value. The low I_s of an ordinary silicon diode might require that R be more than 10^7 ohms , which would require the audio amplifier to have an inconveniently high input impedance. The power dissipated in the detector is V_{det}^2/R . This detector would typically produce, say, 2 V (to operate up in the linear range), which corresponds to a power dissipation of $2^2/10^4 = 0.4 \text{ mW}$. For a given signal strength at the receiver input, the RF stages must have enough gain to produce 0.4 mW .

AC-coupled diode detector

Sometimes circuit considerations require that the detector input be ac coupled. In this case a dc return must be furnished for the detector diode. Such a circuit is shown in Figure 18.4 where an RF choke (a large value inductor) provides this

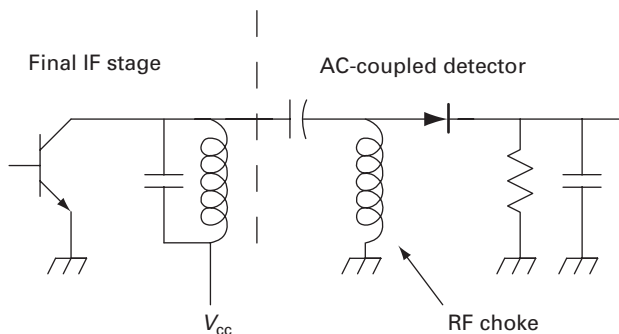


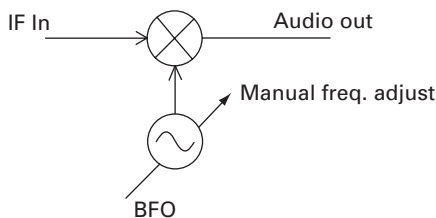
Figure 18.4. AC-coupled diode detector.

dc return path, forcing the average voltage at the left side of the diode to be zero, as in the circuits of Figures 18.1 and 18.2.

18.1.2 Product detectors

Demodulation of SSB and Morse code (cw) signals is usually done by a *product detector* which mixes (multiplies) the IF signal with a locally generated carrier. The resulting difference frequency components become the demodulated signal, while the sum frequency components are discarded. This is shown in Figure 18.5. The free-running oscillator is historically known as a *beat frequency oscillator* (BFO).

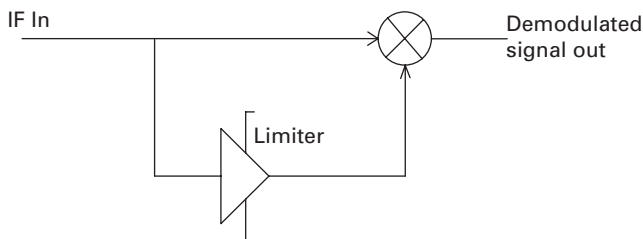
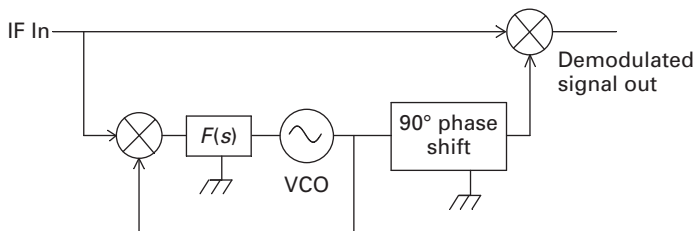
Figure 18.5. Product detector for SSB and CW.



For SSB reception of voice signals, the frequency of the BFO is manually adjusted until the audio sounds approximately natural. For Morse code reception, the BFO is deliberately offset to produce an audible tone, the “beat note.” This was first done by the radio pioneer, Reginald Fessenden. In his heterodyne detector, the predecessor to the Edwin Armstrong’s superheterodyne, the incoming signal voltage was combined with the voltage from a local oscillator,² which was a small arc source, an early negative resistance oscillator.

A product detector can also be used for AM demodulation. Again the signal is *multiplied* by a locally generated carrier. Here the local carrier must have the frequency and phase of the received carrier. Any error in frequency creates a strong audio beat note with the carrier of the received signal and an error in phase reduces the amplitude. Nevertheless, the product detector overcomes the limitations of diode envelope detectors; the input signal levels do not have to be as high and there is no low-signal threshold below which the detector is useless. When the AM signal is consistently strong (usually the case for most broadcast listeners) the local carrier can be a hard-limited version of the input signal. This works because, in double sideband AM, the modulated signal has the same zero crossings as the unmodulated carrier. The product detector shown in Figure 18.6 uses this method. This detector is commonly used for the video detector in analog television receivers.

² In his heterodyne patent of 1902, Fessenden proposed that the transmitting station send *two* signals, closely spaced in frequency. At the receiver, the nonlinear detector would produce an audible beat note. One of these signals could be continuous (not keyed). Instead, it became practical to produce the continuous signal at the receiver site using a *local* oscillator.

Figure 18.6. Product AM detector.**Figure 18.7.** Synchronous AM detector.

The *synchronous detector*, shown in Figure 18.7, is an improved product detector circuit in which a phase lock loop is used to generate the local carrier. The carrier of the input signal provides the reference signal for the loop. In a practical circuit, a limiting amplifier can be used at the reference input to make the loop dynamics independent of the signal level.

The PLL gives the synchronous detector a flywheel effect: the narrowband loop maintains the regenerated carrier during abrupt selective fades (common in short-wave listening). This prevents distortion, common in short-wave receivers, caused by momentary dropouts of the carrier. The PLL provides, in effect, a bandpass filter so narrow that its output cannot change quickly. Note the 90° phase shift network; if the phase detector is the standard multiplier (mixer), the VCO output phase differs from the reference phase by 90°. Without the network to bring the phase back to 0°, the output of the detector would be zero.

18.1.3 Digital demodulation of AM

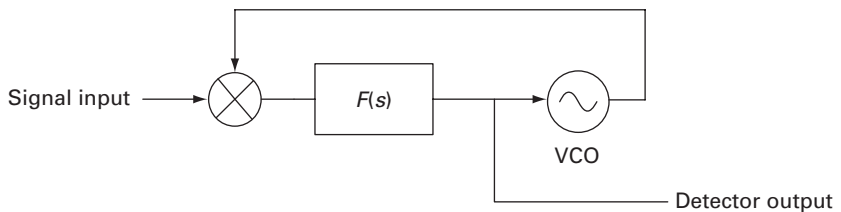
The analog AM demodulators discussed above can be implemented as well in digital circuitry. Envelope detection via full-wave rectification of the IF signal can be done by simply taking the absolute value of the numbers produced by the A-to-D converter. For digital processing, the IF signal is often converted to two baseband signals, I and Q , which result from mixing the IF signal with $\cos(\omega_{IF}t)$ and $\sin(\omega_{IF}t)$. In this case, the amplitude of the signal is given by $(I^2 + Q^2)^{1/2}$, which can also be computed digitally, though not as easily as $|V|$. Synchronous detection can be done.

18.2 FM demodulators

A variety of circuits, often called *discriminators*, have been used to demodulate FM. Most of these circuits are sensitive to amplitude variations as well as frequency variations, so the signal is usually amplitude limited before it arrives at the FM detector. This reduces the noise output, since amplitude noise is eliminated (leaving only phase noise). In addition, it ensures that the audio volume is independent of signal strength.

18.2.1 Phase lock loop FM demodulator

We have already pointed out that a phase lock loop may be used as an FM discriminator. As the loop operates, the instantaneous voltage it applies to the voltage-controlled-oscillator (VCO) is determined by the reference frequency, which here is the signal frequency. The linearity of the VCO determines the linearity of this detector, shown in Figure 18.8.

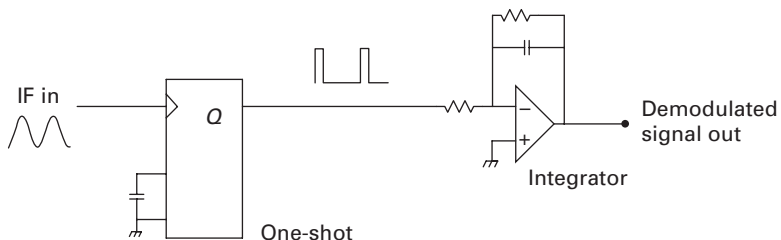


18.8. Phase lock loop FM detector.

18.2.2 Tachometer FM detector

A tachometer FM detector or “pulse counting detector”, shown in Figure 18.9, is just a one-shot multivibrator that fires on the zero crossings of the signal. Each positive zero crossing produces a constant-width output pulse. The duty cycle of the one-shot output therefore varies linearly with input frequency so, by integrating the output of the one-shot, we get an output voltage that varies linearly with frequency.

Figure 18.9. Tachometer FM detector.



18.2.3 Delay line FM detector

The *delay line discriminator* is often used in C-band satellite television receivers to demodulate the FM-modulated video and sound. The IF frequency in these receivers is typically 70 MHz. Figure 18.10 shows a quarter-wave delay line (which could be a piece of ordinary transmission line) which delays the signal at one input of the multiplier. If the input signal is $\cos((\omega_0 + \delta\omega)t)$, then the signal at the output of the delay line is $\cos((\omega_0 + \delta\omega)(t - \tau)) = \sin((\omega_0 + \delta\omega)t - \tau\delta\omega)$ and the baseband component at the output of the multiplier is $-\sin(\tau\delta\omega)$. For small $\tau\delta\omega$ this is just $-\tau\delta\omega$. The output voltage is thus proportional to the frequency offset, $\delta\omega$. If the delay line is lengthened by an integral number of half-wave lengths, the sensitivity of the detector is increased, i.e., a given shift from center frequency produces a greater output voltage.

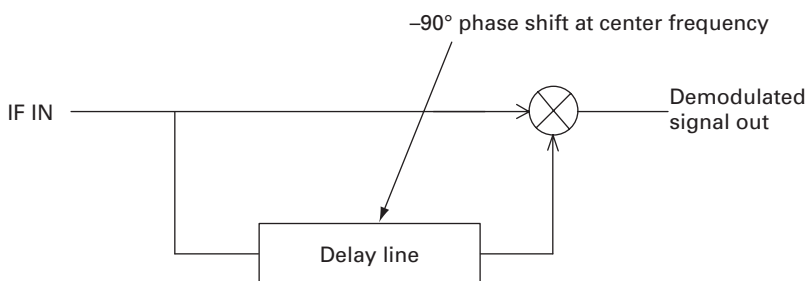
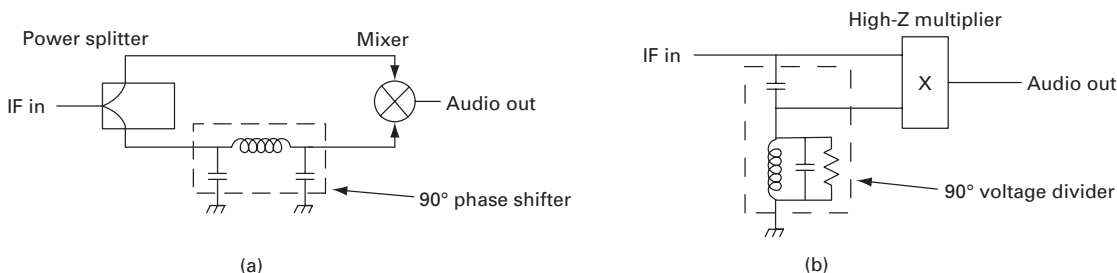


Figure 18.10. Delay line FM detector.

FM quadrature demodulator

The *quadrature FM demodulator*, shown in Figure 18.11, is the same as the delay line discriminator except that an *LC* network is used to provide the delay, i.e., a phase shift that varies linearly with frequency. These circuits are commonly used in integrated circuits for FM radios and television sound; the *LC* networks or an equivalent resonator is normally an off-chip component.

Figure 18.11. Quadrature FM detectors. (a) An *LC* network is used as the delay element. The multiplier provides the necessary resistive termination, (b) A voltage divider provides the necessary phase shift (see Problem 18.3).



18.2.4 Slope FM detector

Slope detection, in which FM modulation is converted to AM modulation, is the original method to demodulate FM. The amplitude response of the IF bandpass filter is made to have a constant slope at the nominal signal frequency. An input signal of constant amplitude will then produce an output signal whose amplitude depends linearly on frequency. A simple envelope detector can detect this amplitude variation. An AM receiver can slope-detect an FM signal if detuned slightly to put the FM signal on the upper or lower sloping skirt of the IF passband filter. A refined balanced slope detector (Figure 18.12) uses two filters with equal but opposite slopes. The filter outputs are individually envelope-detected and the detector voltages are subtracted. This makes the output voltage zero when the input signal is on center frequency, f_0 , and also linearizes the detector by cancelling even-order curvature, such as an $(f-f_0)^2$ term, in the filter shape.

The Foster–Seeley “discriminator,” a classic FM detector, is an example of a balanced slope detector, though this is hardly obvious from the circuit, shown in Figure 18.13.

Note that the transformer has a capacitor across both its primary and its secondary. If the transformer had unity coupling, a single capacitor on one side or the other would suffice. The fact that there are two capacitors tells us the

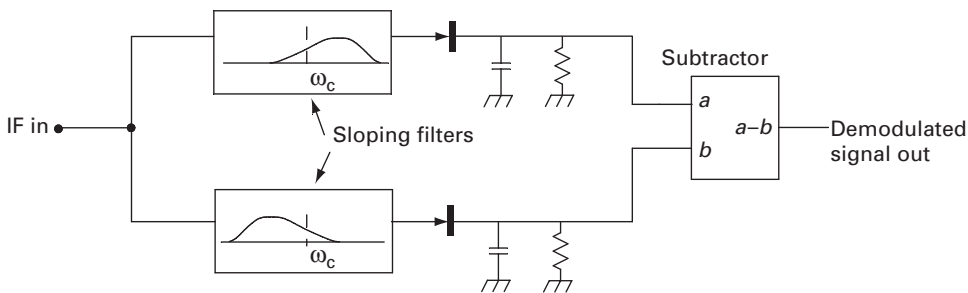


Figure 18.12. Balanced slope FM detector.

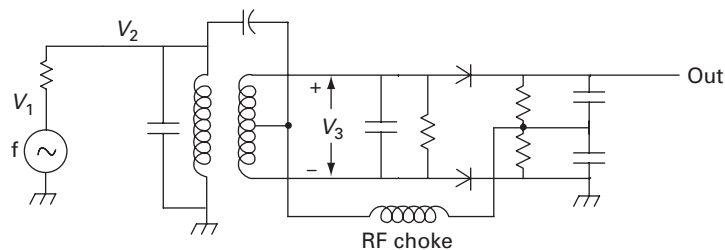


Figure 18.13. Foster–Seeley FM detector.

coupling is not unity; leakage inductance is an element in this circuit. In fact, the leakage inductance, and the magnetizing inductance, are used in a phase shift network, shown in Figure 18.14(a), used to produce the FM-to-AM conversion.

In this circuit, L_1 and L_2 are, respectively, the leakage and magnetizing inductances of the transformer. At the center frequency, f_c , V_2 lags V_3 by 90° . As the frequency increases from f_c , this lag increases. Mostly because of the change in relative phase over the operating region, the magnitude of the vector sum $V_2 + 0.5V_3$ decreases with frequency, while the magnitude of the vector sum $V_2 - 0.5V_3$ increases with frequency. These two combinations correspond to the outputs of the sloping filters in Figure 18.12.

Figure 18.15 is an equivalent circuit of the Foster–Seeley detector using an equivalent circuit for a transformer made up of the leakage inductance, the magnetizing inductance, and an ideal transformer. We will assume for convenience that the primary and secondary inductances and the coupling coefficient have values that give the ideal transformer a 1:1 ratio (see Problem 14.9).

Comparing Figures 18.15 and 18.13, you can see how the top and bottom halves of the transformer secondary are used to add $0.5V_3$ and $-0.5V_3$ to V_2 . The

Figure 18.14. (a) Phase shift network: (b) magnitude vs. frequency of $V_2 + 0.5V_3$ and $V_2 - 0.5V_3$.

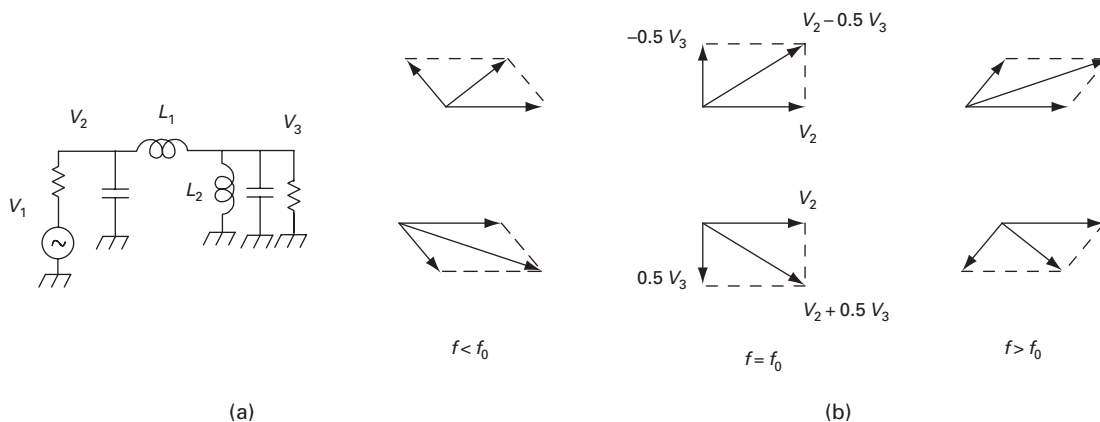
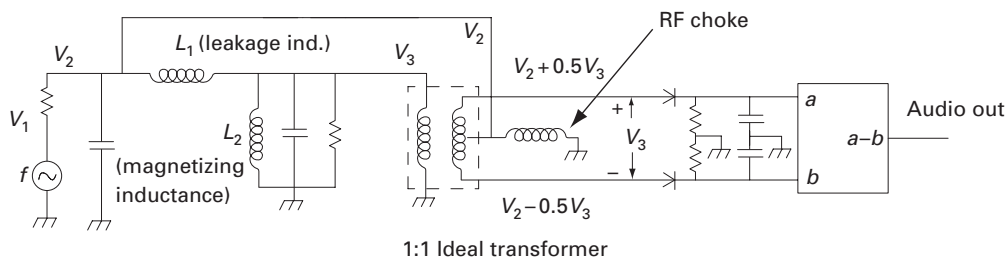


Figure 18.15. Foster–Seeley detector equivalent circuit.

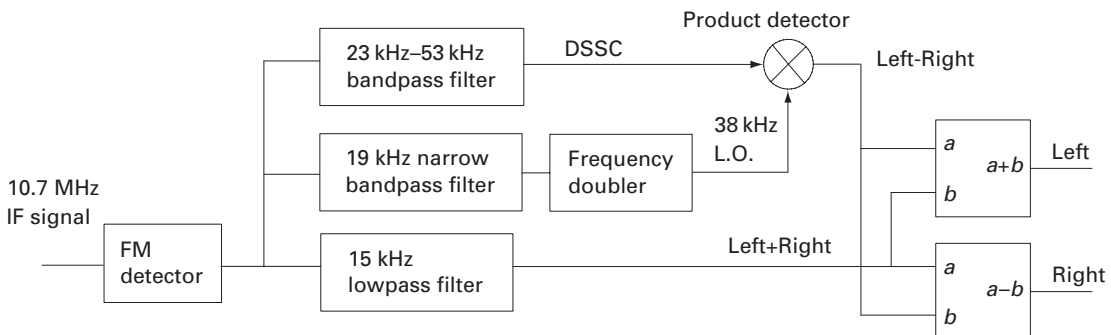


diode detectors produce voltages equal to the magnitudes of $V_2 + 0.5V_3$ and $V_2 - 0.5V_3$ and these magnitude voltages are subtracted, as in Figure 18.12, to produce the output. The RF choke provides a dc return for the diodes. Note that the subtractor can be eliminated by moving the grounding point in the secondary circuit to the bottom of the transformer's secondary winding. The dc blocking capacitor that bridges the transformer is only needed if there is a dc voltage on the primary winding, as when the signal source is the collector of a transistor biased through the primary winding. Finally, note that the capacitor and resistor in parallel with the magnetizing inductance L_2 in the equivalent circuit are actually located on the secondary side of the transformer which is equivalent to being on the left-hand side of the 1:1 ideal transformer.

18.2.5 FM stereo demodulator

Stereo FM, a compatible add-on dating to the 1960s, transmits an L+R (left plus right) audio signal in the normal fashion, and this signal is used by monaural receivers (or stereo receivers switched to “mono”). At the transmitter, the L–R audio is multiplied by a 38 kHz sine wave to produce a DSBSC (double side-band suppressed carrier) signal. This signal, well above the audio range, is added to the L+R audio signal, together with a weak 19 kHz sine wave “pilot” signal. The sum of these three signals, an example of *frequency division multiplexing* (FDM), drives the VCO (or equivalent) to produce the FM signal. At the receiver, the sum signal is demodulated by any ordinary FM demodulator. After demodulation, the L–R signal is brought back down to baseband by a product detector, i.e., a multiplier with a 38 kHz L.O. This L.O., which must have the correct phase (and therefore also the correct frequency), is derived by putting the pilot signal through a frequency doubler. This is shown in the block diagram of Figure 18.16.

Figure 18.16. FM stereo broadcast receiver broadcasting block diagram.



Note that the L–R signal has been doubly demodulated: first by the FM detector and then by the AM product detector. In this stereo system, the L–R signal is more susceptible to noise than the L+R signal. For “full quieting” stereo reception, about 20 dB more signal strength is required than for equivalent monaural reception. For this reason, FM stereo receivers are provided with manual or automatic switches to select monaural (L+R only) operation.

18.2.6 Digital demodulation of FM

When the IF signal has been converted into baseband I and Q signals, the phase of the IF sample is given by $\tan(\theta(t)) = Q(t)/I(t)$. Since frequency is the time derivative of phase, we take the differential of this expression: $(1 + \tan^2 \theta) \delta\theta = \delta Q/I - Q\delta I/I^2$. Substituting Q/I for $\tan \theta$ yields

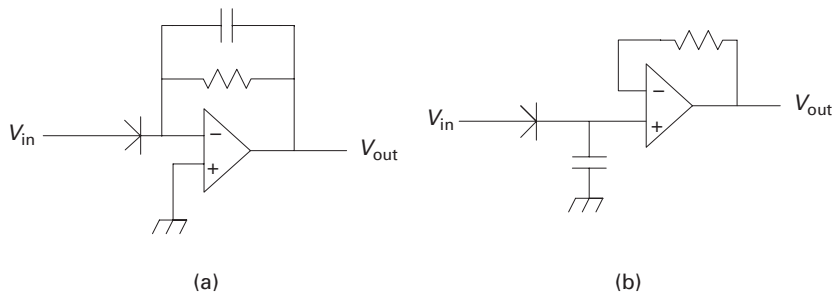
$$\delta\theta = \frac{I\delta Q - Q\delta I}{I^2 + Q^2}. \quad (18.5)$$

For each sampled I, Q pair, $\delta\theta$ is calculated from this pair and the previous pair. The resulting values of $\delta\theta$ are proportional to $d\theta/dt = \omega$. (Note that this could also be done with analog circuitry. Only two analog multipliers would be needed if the IF signal has passed through a limiter, making $(I^2 + Q^2)$ a constant.)

18.3 Power detectors

Square-law detectors are not used as demodulators but are used in laboratory instruments that measure power (wattmeters and rms voltmeters). If we are measuring a sine wave we know that the rms voltage is equal to the peak voltage divided by $\sqrt{2}$. When we know the shape of the waveform, a true square-law meter is not necessary. Even noise power can sometimes be estimated with other than a square-law device. The power of a Gaussian random noise source can be measured by averaging the output of a V^{2n} law device or a $|V|^n$ law device where $n = 1, 2, \dots$. The square-law device, however, is always the optimum detector in that it provides the most accurate power estimate for a fixed averaging time. When we need the optimum power measuring strategy or if we need to measure the rms voltage of an unknown waveform, we must average the output from a true square-law device. “True rms voltmeters” built for this purpose use a variety of techniques to form the square of the input voltage. Some instruments use a network of diodes and resistors to form a piecewise approximation of a square-law transfer function. Other instruments use a thermal method where the unknown voltage heats a resistor. The temperature of the resistor is monitored by a thermistor while a servo circuit removes or adds dc (or sine wave ac) current to the resistor to keep it at a constant temperature. The diode network

Figure 18.17. (a) Simple diode power detector; (b) preferred detector.



requires large signals and the thermal method has a very slow response. Gilbert cell analog multipliers can be used to square the input voltage, but they are limited to relatively low-frequency signals. Generally, when the power is very low and/or the frequency is very high, and/or a very wide dynamic range is needed, the square-law detector uses a semiconductor diode. A simple (but not particularly recommended) circuit is shown in Figure 18.17. In this circuit the average current through the diode is converted to a voltage by the op-amp. The virtual ground at the op-amp input ensures that V_{in} is applied in full to the diode.

It is important to remember that the I vs. V curve for a diode does not follow a square law, even in a limited region. Rather, the law of the diode junction is exponential: $I = I_s[\exp(V/V_T) - 1]$ where V_T , the thermal voltage, is 26 mV. The diode is used at voltages much smaller than V_T so it is permissible to write $I = I_s[V/V_T + 1/2(V/V_T)^2 + 1/6(V/V_T)^3 + 1/24(V/V_T)^4]$. Obviously we can restrict the input voltage enough to neglect the last term. But what about the first term, the linear term? This dominant term will provide a current component that has the same frequency spectrum as the input signal; if that spectrum extends down to zero the output of the detector will be corrupted with extra noise. The third-order term will also have a baseband component. But if the input signal is a bandpass signal, the first-order and third-order terms are high-frequency signals that will be eliminated by whatever lowpass filter is applied to do the averaging (the capacitor in the above circuit). The simple square-law diode detector, then, is appropriate for measuring signals whose frequency components do not extend down into the baseband output spectrum of the detector circuit. (You can think of various two-diode balanced circuits to cancel the linear component but remember that the two diodes must be matched very closely to start with and then maintained at the same temperature.) The diode and op-amp circuit shown above serves to explain why diode detectors are used for bandpass signals but not for baseband signals. A better circuit – the preferred circuit – is shown in Figure 18.17(b). Here the sensitivity is not dependent on I_s and the circuit has a much better temperature coefficient. In this circuit no dc current can flow in the diode. The capacitor, however, ensures that the full ac signal voltage is applied to the diode. Expanding the exponential relation for the diode current and taking only the dc components we have

$$0 = \frac{-V_{dc}}{V_T} + \frac{\frac{1}{2}\langle V_{in}^2 \rangle}{V_T^2} \quad (18.6)$$

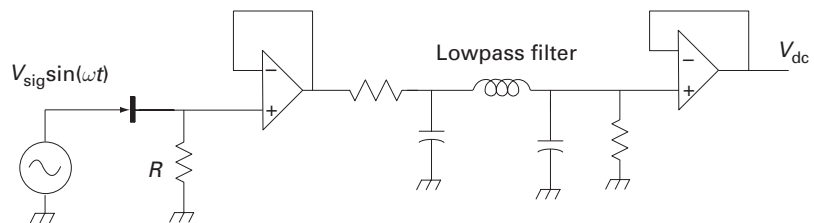
and

$$V_{dc} = \frac{\frac{1}{2}\langle V_{in}^2 \rangle}{V_T} \quad (18.7)$$

which shows that the dc output voltage is indeed proportional to the average square of the input voltage.

Problems

Problem 18.1. Assume that in the envelope detector circuit shown below, the diode is a perfect rectifier (zero forward resistance and infinite reverse resistance) and that the op-amps are ideal.



- (a) Calculate the efficiency, V_{dc}/V .
- (b) Calculate the effective load presented to the generator (the value of a resistor that would draw the same average power from the source).

Problem 18.2. Draw a vector diagram to show why an envelope detector will produce an audio tone when it is fed with the sum of an IF signal and a (much stronger) BFO signal.

Problem 18.3. Show that the voltage divider network in Figure 18.11(b) can produce an output voltage shifted 90° from the input voltage when the input voltage is at a specified center frequency, ω_0 . Determine the position of the output phase when the input signal is slightly higher or slightly lower than ω_0 .

Problem 18.4. When an interfering AM station is close in frequency to a desired AM station, an audio tone “beat note” is produced, no matter whether the receiver uses an envelope detector or a product detector. (In the case of an envelope detector, the beat note is produced because the amplitude of the vector sum of the two carriers is effectively modulated by an audio envelope. In the case of the product detector, the carrier of the undesired station acts as a modulation sideband and beats with the BFO.) Will the same thing happen with FM? Suppose two carriers (i.e., cw signals), separated by say, 1 kHz, appear in the IF passband of an FM receiver. Let their amplitudes be in the ratio of, say,

1:10. Draw a phasor diagram of the sum of these two signals. Does the vector sum have phase modulation? Will the receiver produce an audio tone? What happens in the case when the amplitudes of the two signals are equal?

Problem 18.5. Try the following experiment with two FM receivers. Tune one receiver to a moderately strong station near the low-frequency end of the band. Use the other receiver (the local oscillator) as a signal generator. (This receiver must have continuous rather than digital tuning.) Turn its volume down and hold it close to the first receiver so there will be local oscillator pickup. Carefully tune the second receiver 10.7 MHz higher in frequency until an effect is produced in the sound from the first receiver. What is the effect? Can you use this experiment to confirm your answer to Problem 18.4?

References

- [1] Gosling, W., *Radio Receivers*, London: Peter Peregrinus, 1986.
- [2] Landee, R. W., Davis, D. C. and Albrecht, A. P., *Electronics Designers' Handbook*, New York: McGraw-Hill, 1957.
- [3] Rohde, U. and Whitaker, J., *Communication Receivers – DSP, Software Radios, and Design*, 3rd edn, New York: McGraw Hill, 2000.