431/531 Class Notes 11

8 Radio Basics

In this section we will discuss some basic concepts concerning signal modulation, generation, receiving, and demodulation. Some of these concepts are quite general and see applications in many areas. However, the most familiar perhaps is that of broadcast radio generation and receiving, hence the title of the section.

We will begin with a simplified discussion of amplitude modulation (AM). From this, we can see how to carry over many of the concepts to other forms of signal modulation and reception of signals.

8.1 The Case for Modulation

Consider the familiar example of radio signals which carry audio information. The audio itself has a typical frequency range of

$$20\text{Hz} < f_{\text{audio}} < 20\text{kHz}$$

Hence, audio has an effective *bandwidth* of about 20 kHz. Even if it were possible to broadcast signals of such low frequency in the electromagnetic spectrum, there would be a multitude of confusion resulting from the interference between competing broadcasts.

On the other hand, electromagnetic signals in the radio-frequency (RF) range, have frequencies roughly from several hundred kHz to several hundred MHz. An audio signal which modulates an RF "carrier" of, say, 1 Mhz, uses only the range 20.00 ± 0.02 MHz. Another broadcast "channel" with a carrier frequency only 100 kHz removed will have give interference with its own signal at 20.10 ± 0.02 MHz.

We will look at several techniques for signal modulation, beginning with amplitude modulation. It is important to remember that the signals do not have to be audio, that is only a familiar example. They could, represent any information which can be converted to an electromagnetic signal. Another familiar example is the modulation of computer-generated signals for transmission over telephone lines.

8.2 Amplitude Modulation

Figure 43 gives the general scheme. Each frequency which is to carry information, $\omega_m = 2\pi f_m$, is "mixed" with the high-frequency carrier frequency, $\omega_c = 2\pi f_c$, to produce an output signal of the form

$$V_s(t) = A\,[1 + m\cos\omega_m t]\cos\omega_c t \tag{45}$$

where A is a constant and the constant $m \leq 1$ is known as the modulation index. We see that the carrier amplitude $A\cos(\omega_c t)$ is modulated by the factor $1+m\cos(\omega_m t)$, where $m = 0$ represents the limit of no modulation and $m = 1$ is a miximally modulated signal.

By using the identity

$$\cos x \cos y = \frac{1}{2}\left[\cos(x+y) + \cos(x-y)\right]$$

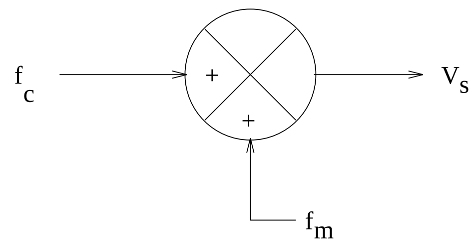

Figure 43: Schematic of modulation.

we can do a "poor man's" Fourier transform of V_s:

$$V_s(t) = A\cos\omega_c t + \frac{1}{2}Am\left[\cos\left((\omega_c+\omega_m)t\right) + \cos\left((\omega_c-\omega_m)t\right)\right] \quad (46)$$

So we have a central carrier frequency plus two side-bands at $f_c \pm f_m$.

8.3 Detection of AM

8.3.1 Heterodyne Detection

We first consider the simple, but subtle, radio receiver shown in Fig. 44. A real receiver might include at the input an antenna followed by an LC bandpass filter, with tunable capacitor. The filter is a resonant circuit with a sharp peak at the carrier frequency of the broadcast $\omega_c = 1/\sqrt{LC}$. The Q of the filter is set so that the width of the peak of the transfer function matches the bandwidth $\Delta\omega$ of the modulating signal, roughly from $\omega_c - \omega_m$ to $\omega_c + \omega_m$. With this addition, and without the amplified output, the passive "crystal" radio receiver looks like this.

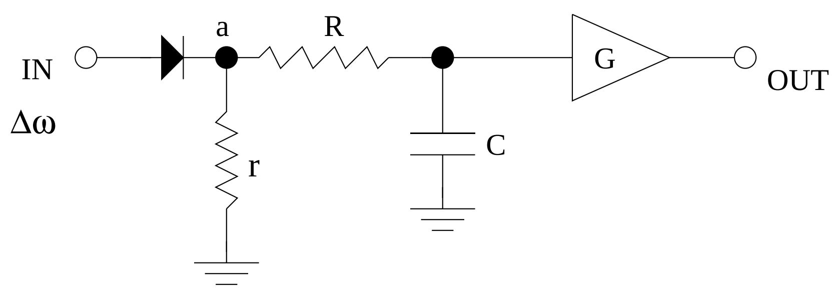

Figure 44: Simple AM receiver.

The resistor R and capacitor clearly form a low-pass filter. The cutoff frequency would be set between ω_m and ω_c in order to keep the information encoded by the low-frequency modulations, and remove the carrier. However, without the diode, the effect would be to throw away all of the information, too, since as we saw from Eqn. 46, all of the frequencies of interest are actually in a narrow band cenetered about the carrier frequency. Without the diode, the system is linear, and no signal will be present at the output.

The diode is non-linear; recall its V-I curve. In order to illustrate how this works, we assume a specific form for the response of a forward-biased diode as $I = bV^2$, where b is a constant. A resistor r is inserted between point a and ground (Fig. 44) in order to convert this diode current to a voltage to be presented to the low-pass filter. Now let V be the linear combination of two signals: $V = V_1 \cos\omega_1 t + V_2 \cos\omega_2 t$. This then gives rise to an output current

$$I = bV_1^2 \cos^2\omega_1 t + bV_2^2 \cos^2\omega_2 t + 2bV_1V_2 \cos\omega_1 t \cos\omega_2 t$$

Again using trigonometric identities to form the poor man's Fourier transform, this becomes

$$2I/b = V_1^2 + V_2^2 + V_1^2 \cos 2\omega_1 t + V_2^2 \cos 2\omega_2 t + 2V_1V_2\left[\cos((\omega_1+\omega_2)t) + \cos((\omega_1-\omega_2)t)\right]$$

Therefore, from the original two frequencies, the diode has produced harmonics (twice the original), as well as the sum and difference.

In the case of our simplified AM broadcast signal of Eqn. 46, where three frequencies are originally present (ω_c and $\omega_c \pm \omega_m$), the effect of the diode is easily generalized from the steps above. We find that the output of the diode will include DC, the first harmonics of all three frequecies, as well as the six possible sum and difference frequencies. Of particular interest for our receiver is the difference frequency between the carrier and the modulated carrier. For example,

$$\omega_c - (\omega_c - \omega_m) = \omega_m$$

Therefore, we do in fact recover a Fourier component corresponding to our original modulating signal. This can then be separated from the higher frequencies using the low-pass filter and amplified. This represents a simple example of so-called *heterodyne detection*, in which different frequencies are combined in order to extract a difference frequency.

It is also interesting to note that with our example $I = bV^2$, we have squared the input. When we examine this is frequency domain (Fourier transform) and low-pass filter the result (averaging), we have effectively formed the so-called *power spectrum* of the input, also called the *spectral power density*.

8.3.2 Harmonic Distortion

Note that we intentionally introduced a non-linear element (the diode) to our system. An unintentional non-linearity in a circuit, for example in an audio amplifier circuit, can introduce additional frequencies as demonstrated above. In particular, our diode with the $I = bV^2$ behavior introduced first harmonics of the original frequencies at twice the original. In general, a non-linearity may include any number of higher-order terms: $I = b_1V + b_2V^2 + b_3V^3 + \cdots$, where each additional power can generate the next higher harmonic. For example, a non-zero b_3 will generate a 2nd harmonic of the original ω at 3ω. The introduction of harmonics of the input signal is called *harmonic distortion*. Since the pattern of harmonics is what distinguishes musical instrument types to the ear, the introduction of non-linearities should be avoided in high-fidelity amplifiers.

8.3.3 Homodyne Detection

An example of this technique is given in the text, pages 653 and 889. It uses a phase-locked loop (PLL) circuit at the input of the receiver. Recall that the PLL circuit is designed to produce an output which is proportional to shifts in phase of the input. Since one can consider the modulation of the carrier to be a phase shift (by amount $\omega_m t$), the output of the PLL can then produce a voltage signal proportional to these phase shifts, which in turn is used to provide active rectification of the input at the frequency of the modulation. The essential non-linear behavior of the diode discussed above is provided in this case by an active voltage multiplier. This type of PLL circuit is actually more relevant to FM detection, discussed below.

8.3.4 Superheterodyne Detection

This technique is illustrated in the text, pages 895-6. This is essentially a fancy version of our simple heterodyne detector above. In this case, the simple passive LC bandpass filter at the input is replaced by a local oscillator and mixer. An example is given in Figure 13.41 of the text. Consider an input carrier of frequency 10 MHz which has amplitude modulated at some much lower frequency. This signal is mixed with a local oscillator of fixed frequency greater that the carrier. In the example of the text, the local oscillator has frequency tuned to be $f_{\rm LO} = 10.455$ MHz, exactly 455 kHz greater than the carrier. As with our earlier diode example, the mixed signal includes the difference frequency, in this case 455 kHz, which in turn has nearby sideband frequencies which differ from 455 kHz by the audio modulation frequencies. From this point on, the detection is carried out as in the simple heterodyne example. One advantage here is that a relatively high-frequency carrier, which in general will be difficult to condition using conventional electronics is effectively reduced to a more manageable frequency, in the example from 10 MHz to 455 kHz. The other advantage is that the band-pass tuning which follows the mixer is always centered at a constant 455 kHz. So the tuning is accomplished by adjusting the oscillator, rather than the filter.

8.4 Other Modulation Schemes

Recall from Eqn. 45 that for AM the amplitude is modulated by varying the frequency ω_m. However, to preserve the information, the generation and receipt of the amplitude must be linear. In addition, most noise sources will naturally appear as voltages, and hence will add to the AM signal. On the other hand, phase and frequency modulation (FM) do not suffer from these complications. Hence, where fidelity is important, these schemes have intrinsic advantages. Radio broadcast by FM also has the additional advantage, by dint of historical accident, of occupying a higher frequency band, thus allowing easy accomodation of a full audio bandwidth. However, unlike the AM radio band, the FM band signals do not reflect from the ionosphere, and therefore can not be transmitted over very large distances (at night).

8.4.1 Phase Modulation

A carrier of frequency ω_c is *phase modulated* if the resulting signal has the form

$$V(t) = V_0 \cos(\omega_c t + A_p \cos\omega_m t) \tag{47}$$

where V_0 and A_p are constants and ω_m is the modulating frequency, as before. This can be expanded, and for $A_p \ll 1$ can also be simplified:

$$\begin{aligned} V(t)/V_0 &= \cos\omega_c t\ \cos(A_p\cos\omega_m t) - \sin\omega_c t\ \sin(A_p\cos\omega_m t) \\ &\approx \cos\omega_c t - A_p \sin\omega_c t\ \cos\omega_m t \end{aligned}$$

$$= \cos\omega_c t - \frac{1}{2}A_p\left[\sin((\omega_c+\omega_m)t) + \sin((\omega_c-\omega_m)t)\right] \qquad (48)$$

As for AM, two new sidebands have appeared, but now they are 90° out of phase with respect to the carrier.

8.4.2 Frequency Modulation

The phase of a sinusoidal function, when frequency is a function of time, can in general be expressed as

$$\phi = \int \omega dt$$

Now suppose the frequency is modulated by a frequency ω about some central carrier frequency

$$\omega = \omega_c + A_f \cos\omega_m t$$

where A_f is a constant. Then the phase becomes

$$\phi = \omega_c t + \frac{A_f}{\omega_m}\sin\omega_m t$$

Here, A_f is called the *frequency deviation* and A_f/ω_m is the *modulation index* for FM. Carrying out steps analogous to those for Eqn. 48 gives the following expression for the FM signal:

$$V(t)/V_0 = \cos\phi = \cos\omega_c t + \frac{A_f}{2\omega_m}\left[\cos((\omega_c+\omega_m)t) - \cos((\omega_c-\omega_m)t)\right] \qquad (49)$$

So again the Fourier spectrum is similar to what we found for AM, except now one of the two sidebands has amplitude of opposite sign.

8.5 FM Detection

In the AM detection schemes discussed above, the diode or other non-linear element is used to extract an output signal proportional to $\cos\omega_m t$, and hence provide a reproduction of the original modulation, for example in the form of an audio signal. For FM detection we need to replace the diode with something which can provide a voltage output proportional to the input frequency modulated signal. We explored such a technique in Lab 5 in the form of the phase-locked loop circuit. The PLL scheme is reproduced in Fig. 45. (In this application, the counter can be omitted.) Recall that the signal before the VCO, labelled V_{out}, is proportional to input phase shifts. This is exactly what we need to detect the phase shift introduced by FM. All that is left is to feed V_{out} to a low-pass filter and amplifier, as before.

An apparent practical limitation of this technique for FM radio reception is that PLLs do not operate at these high frequencies ($\sim$ 100 MHz). This is overcome by using the technique discussed above at the front-end of the superheterodyne receiver. The input is mixed using a local oscillator and the resulting lower frequency (455 kHz in our example) modulated signal is then input to the PLL. Another technique, called quadrature detection is briefly discusses in the text, page 652.

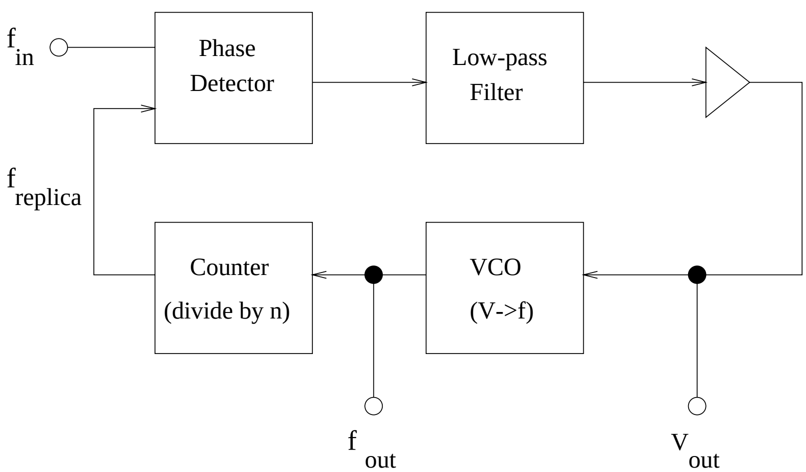

Figure 45: PLL schematic. V_{out} provides the FM signal.