3 Frequency Modulation

Frequency Modulation (FM) is a type of angle modulation. In angle modulations, unlike amplitude modulations, there is no simple relation beteen the bandwidths of the modulating and modulated signals. Moreover, the same modulating signal renders angle-modulated signals with different bandwidths depending on certain parameters of the modulator. Indeed, one of the main advantages of angle modulations, especially FM, is that the quality (signal-to-noise ratio) of the demodulated signal can be improved by increasing the bandwidth of the transmitted signal. Another great advantage of angle modulations is that they are robust against amplitude distortions (saturation, amplitude clipping, variable attenuation,...) because, as we will discuss later, the information is embedded in the instantaneous phase of the modulated signal. Among its disadvantages with respect to AM (DSB-LC) we can mention a smaller reception range and a greater complexity of modulators and demodulators.

In this assignment, we briefly describe the fundamental principles of angle modulation, making special emphasis on FM because of its relevance. For a more detailed description, students are referred to the bibliography.

3.1 Angle Modulation

3.1.1 Basic concepts [3][2]

As we already know, *modulation* is the process of modifying a given parameter of a carrier signal (amplitude, frequency or phase) according to the variations of a modulating signal or message. Thus, considering a carrier

$$c(t) = A_c \cos(2\pi f_c t + \phi_c), \tag{3.1}$$

we will obtain an angle-modulated signal (either in phase or frequency) when the phase, ϕ_c, varies in time according to the modulating signal. In general, we can write the expression of an angle-modulated signal as

$$x(t) = A_c \cos(2\pi f_c t + \phi(t)), \tag{3.2}$$

where the argument of the cosine function $\theta_i(t) = 2\pi f_c t + \phi(t)$ is the *instantaneous phase* and $\phi(t)$ the *instantaneous phase deviation.*

The *instantaneous frequency*, $f_i(t)$, is simply the derivative with respect to time of the instantaneous phase (divided by 2π to obtain the result in Hz), i.e,

$$f_i(t) = \frac{1}{2\pi}\frac{d\theta_i(t)}{dt} = f_c + \frac{1}{2\pi}\frac{d\phi(t)}{dt}, \tag{3.3}$$

where the term $\frac{1}{2\pi}\frac{d\phi(t)}{dt}$ is the *instantaneous frequency deviation.*

Angle modulation reduces to varying the instantaneous phase or frequency deviation according to the modulating signal, $m(t)$. In the first case (phase modulation, PM) we make

$$\phi(t) = k_p m(t), \tag{3.4}$$

where k_p is the so-called *phase deviation constant* or *phase sensitivity* with rad/V as units. Hence, the expression of the resulting modulated signal is

$$x_{\mathrm{PM}}(t) = A_c \cos(2\pi f_c t + k_p m(t)), \tag{3.5}$$

and, of course, the *maximum phase deviation* is given by $\Delta\phi_{max} = k_p \max|m(t)|$ (rad).

As already mentioned, in the case of *frequency modulation* (FM) the variations of the modulating signal alter the instantaneous frequency deviation, thus

$$\frac{1}{2\pi}\frac{d\phi(t)}{dt} = k_f m(t) \tag{3.6}$$

where k_f is the *frequency deviation constant* or *frequency sensitivity* (units Hz/V). Then it follows that

$$\phi(t) = 2\pi k_f \int_{-\infty}^{t} m(u)du, \tag{3.7}$$

$$x_{\mathrm{FM}}(t) = A_c \cos\left(2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(u)du\right), \tag{3.8}$$

where the *maximum frequency deviation* is $\Delta f_{max} = k_f \max|m(t)|$.

Comparing expressions (3.5) and (3.8) we immediately realize that an FM modulator is equivalent to an integrator followed by a PM modulator; and that a PM modulator is equivalent to a differentiator followed by an FM modulator.

3.1.2 Frequency Modulation in Matlab

The code in `toneFMmod.m` shows an example of FM modulation, where the modulating signal is a sinusoid.

toneFMmod.m

```
fs = 32000;                              % Sampling frequency (Hz)
kf = 500;                                % Frequency deviation constant (Hz/V)
fc = 3000;                               % Carrier frequency (Hz)
fm = 200;                                % Frequency of modulating signal (Hz);
N = 10000;                               % Number of samples
n = 0:N-1;
Ts = 1/fs;
t = n*Ts;                                % Sampling times
m = cos(2*pi*fm*t);                      % Modulating signal
x = cos(2*pi*fc*t+2*pi*kf*cumsum(m)*Ts); % FM signal
```

□ **Question 1**

Execute the code in `toneFMmod.m`.

a) In a single figure, plot the first 500 samples of the modulating and modulated signals as a function of time. Notice the changes in the plot as k_f varies in the interval $[200, 800]$ Hz/V.

b) Using `powerspec` plot the power spectral density of the modulated signal and observe how it changes when modifying k_f and/or f_m. For $k_f = 250$ Hz/V and $f_m = 200$ Hz, write down the frequency location of the most relevant spectral components. Is there any relation among them?

c) Repeat the two previous points, but now considering as modulating signal a rectangular wave varying between $+1$ and -1 with the same frequency as the previous modulating sine wave. ■

3.1.3 Spectral characteristics of angle-modulated signals [3]

The non-linear character of angle modulation makes its analytical study quite complicated, if not impossible, except for cases where the modulating signal is extremely simple. One of these simple cases takes place when the modulating signal is a sinusoid, which is briefly described below.

Angle modulation by a sinusoid

When the modulating wave is a tone (sinusoid) with frequency f_m , the PM or FM modulated signal is given by

$$\begin{aligned} x(t) &= A_c \cos(2\pi f_c t + \beta \cos(2\pi f_m t)) \quad (3.9)\\ &= A_c \cos(\omega_c t + \beta \cos(\omega_m t)), \quad (3.10) \end{aligned}$$

where the *modulation index* β is defined as

$$\beta = \begin{cases} k_p \max|m(t)| & \text{in PM,} \\ \frac{k_f \max|m(t)|}{f_m} = \frac{\Delta f_{\max}}{f_m} & \text{in FM.} \end{cases} \quad (3.11)$$

In this case, and making use of Fourier Series analysis, it can be shown that the power spectral density (PSD) of the angle-modulated signal (in phase or frequency) is given by[1]

$$S_X(f) = \frac{A_c^2}{4} \sum_{n=-\infty}^{\infty} J_n^2(\beta) \left[\delta(f - f_c - nf_m) + \delta(f + f_c + nf_m)\right], \quad (3.12)$$

where $J_n(\beta)$ denotes the Bessel function of the first kind, of order n and argument β.

Note that the number of spectral components in (3.12) is infinite. Fortunately, it tirns out that only a reduced number of them have significant amplitude (see Figures 3.1 and 3.2), and the bandwidth of the modulated signal can be estimated by the following rule of thumb:

$$B_{\mathrm{FM}} = 2(\beta + 1) f_m, \quad (3.13)$$

so that the estimated bandwidth B_{FM} approximately contains 98% of the power of the modulated signal, P_x. Moreover, since that $\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1$ for any β, then

$$P_x = \int_{-\infty}^{\infty} S_X(f) df = \frac{A_c^2}{2}, \quad (3.14)$$

being equal to the power of the carrier regardless the modulating signal (i.e., (3.14) holds even if $m(t)$ is not a sinusoid).

Expression (3.13) shows that the bandwidth depends on both the modulation index and the frequency of the modulating sinusoid. If $\beta \ll 1$, the resulting FM modulation is said to be *narrowband,* since $B_{\mathrm{FM}} \approx 2f_m$ (similar bandwidth as in AM). Otherwise, the FM modulation is regarded as *wideband.*

[1] In such PSD analysis, an additional phase term ϕ_c is considered in the modulated signal, modeling the uncertainty about the initial phase of the oscillator. As usual, this phase term is modeled as a uniform random variable in the interval $[0, 2\pi)$.

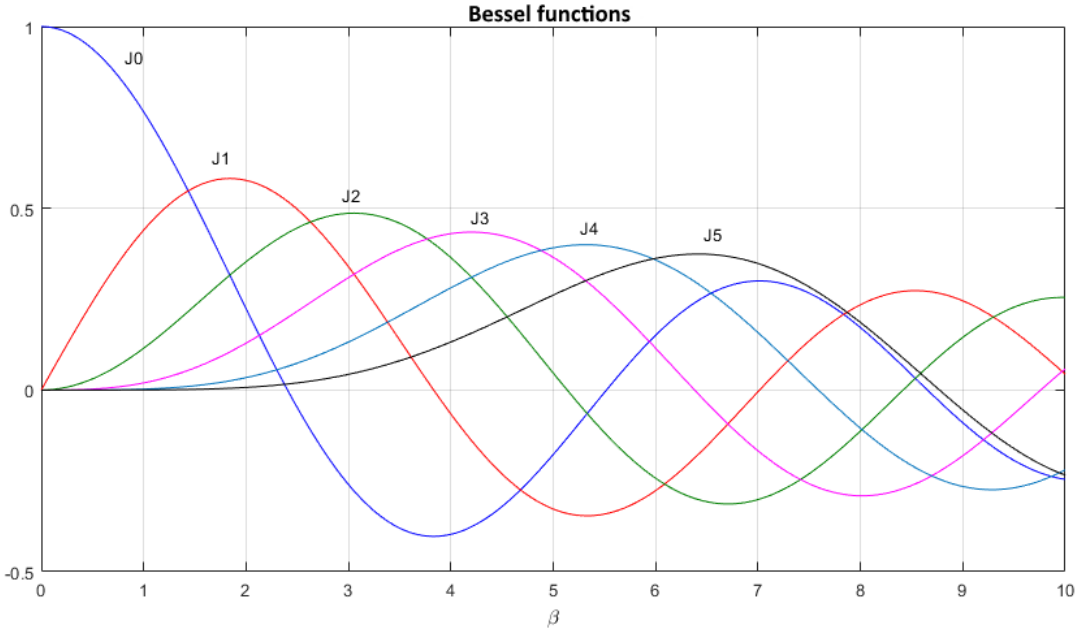

Figure 3.1: Bessel functions of first kind of increasing orders vs argument β

n \ β	0,1	0,2	0,5	1	2	5	8	10
0	0,9975	0,9900	0,9385	0,7652	0,2239	-0,1776	0,1717	-0,2459
1	0,0499	0,0995	0,2423	0,4401	0,5767	-0,3276	0,2346	0,0435
2	0,0012	0,0050	0,0306	0,1149	0,3528	0,0466	-0,1130	0,2546
3		0,0002	0,0026	0,0196	0,1289	0,3648	-0,2911	0,0584
4			0,0002	0,0025	0,0340	0,3912	-0,1054	-0,2196
5				0,0002	0,0070	0,2611	0,1858	-0,2341
6					0,0012	0,1310	0,3376	-0,0145
7					0,0002	0,0534	0,3206	0,2167
8						0,0184	0,2235	0,3179
9						0,0055	0,1263	0,2919
10						0,0015	0,0608	0,2075
11						0,0004	0,0256	0,1231
12						0,0001	0,0096	0,0634
13							0,0033	0,0290
14							0,0010	0,0120
15							0,0003	0,0045
16							0,0001	0,0016

Values not in the table are negligible.

For negative orders,

$$J_{-n}(\beta) = (-1)^n J_n(\beta)$$

Figure 3.2: Table of Bessel functions of the first kind according to their order, n, and argument, β.

□ **Question 2**

Consider again the code in `toneFMmod.m`.

a) For the default parameter values, vary the frequency of the modulating signal and observe the position of the resulting spectral components.

b) What parameters affect the modulation index?

c) Adjust the value of k_f to vary the modulation index between 0.1 and 4. With the help of `powerspec`, plot the power spectral density of the modulated signal and, from the graphs, estimate its bandwidth.

d) For the values of the modulation index considered in the previous point, estimate the bandwidth according to expression (3.13) and compare the results with those estimated from the graphs. ■

Extension to any modulating signal

Obviously, for a generic $m(t)$ with more than one spectral component, the PSD no longer is given by expression (3.12). However, an approximation known as *Carson's rule* can be used to estimate the bandwidth. It is a straightforward generalization of the expression (3.13), obtained by replacing f_m by the bandwidth of the modulating signal:

$$B_{\mathrm{FM}} = 2(D+1)W, \tag{3.15}$$

where W is the bandwidth of the modulating signal and $D = \Delta f_{\max}/W$ is the so-called *deviation ratio* [2].

□ **Question 3**

In `audioFMmod` a FM signal is generated from an audio message. Consider different values of k_f in the range 25 to 200 kHz/V and observe their influence in the bandwidth of the modulated signal. Compare the observed bandwidths to those estimated by Carson's rule. ■

3.2 FM demodulation

There are multiple methods for demodulating an FM signal, ranging from the simplest ones (as for example a zero-crossing counter) to the most complex and recent (Phase Locked Loops, PLLs). As usual, the demodulation methods can be classified as coherent (PLLs, for example) or incoherent. In this practice we will study two incoherent methods.

3.2.1 FM-to-AM conversion plus envelope detection [2]

The structure of this receiver is illustrated in Figure 3.3. Let us momentarily suppose that the limiter is not present, so that the FM signal is directly fed to the differentiator. Given

$$x(t) = A_c \cos\left(2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(u)du\right), \tag{3.16}$$

the output of the differentiator is

$$w(t) = -2\pi A_c\left(f_c + k_f m(t)\right) \sin\left(2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(u)du\right), \tag{3.17}$$

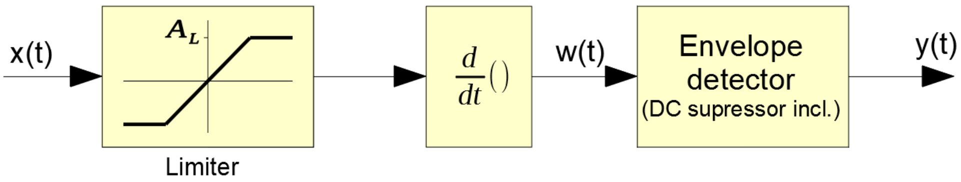

Figure 3.3: FM demodulator with a differentiator and an envelope detector (adapted from [2]).

which is analogous to a DSB-LC modulation, but with a frequency-varying carrier. Assuming that $f_c \gg k_f m(t)$ for all t, we can use an envelope detector to obtain

$$\begin{aligned} y_0(t) &= 2\pi A_c \left| f_c + k_f m(t) \right| && (3.18) \\ &= 2\pi A_c \left(f_c + k_f m(t) \right), && (3.19) \end{aligned}$$

and eliminating the DC component, we get

$$y(t) = 2\pi A_c k_f m(t). \tag{3.20}$$

The purpose of the limiter in Figure 3.3, whose input-output operation is described as

$$s(t) = \begin{cases} -A_L, & \text{if } x(t) < -A_L, \\ x(t), & \text{if } -A_L \leq x(t) \leq A_L, \\ A_L, & \text{if } x(t) > A_L, \end{cases} \tag{3.21}$$

is to ensure that the maximum amplitude of the wave at the differentiator input is constant[2], even at the cost of slightly clipping the FM signal (the message information remain in its instantaneous frequency). When including the limiter the above expressions are approximately valid simply replacing A_c by A_L.

☐ **Question 4**

The code in `envFMdemod` simulates the previous demodulator.

a) Analyze the code and find the influence of k_f on the demodulated signal (consider the range from 25 to 200 kHz/V).

b) In order to study the behaviour of the demodulator in the presence of noise, modify the VN variable to take the values 0.06 and 0.03. Observe the influence of the noise on the PSD of the demodulated signal.

c) Now set the noise power to 0.03 and modify the value of k_f in the range of values in a). How does this parameter influence the perceived quality?

d) The analysis of the behaviour in the presence of noise is simpler if we first observe the effect of the noise in the absence of a received signal. To do so, multiply the FM signal by zero and observe the shape of the noise PSD at the demodulator output. Is it approximately constant across the bandwidth of interest?

e) Can you think of a way to alleviate the effect observed in the previous point?

f) We can improve the performance of the demodulator in Figure 3.3 by introducing an adequate band-pass filter after the limiter. Why?

■

[2] In a real system, the signal at the limiter input contains noise and amplitude fluctuations caused by the channel. The limiter helps to reduce the impact of these deleterious effects on the final demodulated signal.

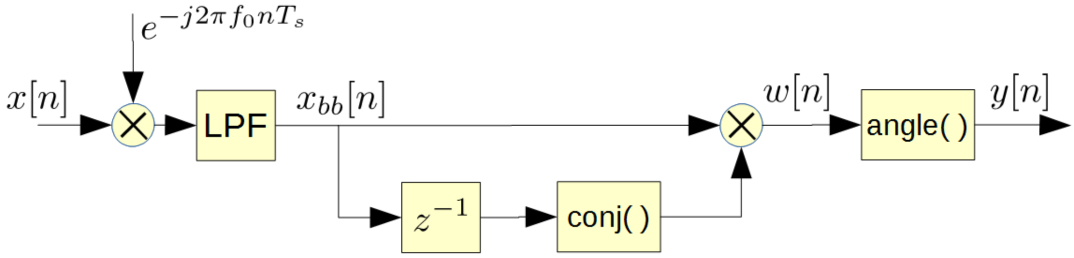

Figure 3.4: Complex Delay Line Discriminator (adapted from [1])

3.2.2 Complex Delay Line Discriminator [1] [4]

Assume that we have an FM signal sampled at an adequate sampling frequency, f_s, obtaining the resulting sequence:

$$x[n] = A_c \cos(2\pi f_c n T_s + \phi[n]) \quad (3.22)$$

$$= A_c \cos(\hat{\omega}_c n + \phi[n]) \quad (3.23)$$

$$= \frac{A_c}{2}\left[e^{j(\hat{\omega}_c n + \phi[n])} + e^{-j(\hat{\omega}_c n + \phi[n])}\right]. \quad (3.24)$$

Suppose now that $x[n]$ is the input to the system shown in Figure 3.4. Multiplying $x[n]$ by a complex exponential with discrete frequency $\hat{\omega}_0 = 2\pi f_0/f_s$ almost identical to $\hat{\omega}_c$ (f_0 very close to f_c in the continuous equivalent), we move the spectral components to frequencies around $\hat{\omega}_0 \pm \hat{\omega}_c$. Then, the low-pass filter will only let through the low-frequency componens, that is, the complex-baseband equivalent sequence, $x_{bb}[n]$, where

$$x_{bb}[n] = \frac{A_c}{2} e^{j(\Delta\hat{\omega}_c n + \phi[n])} \quad (3.25)$$

with $\Delta\hat{\omega}_c = \hat{\omega}_c - \hat{\omega}_0$.

The sequences $w[n]$ and $y[n]$ shown in Figure 3.4 will be given by

$$w[n] = x_{bb}[n] x^*_{bb}[n-1] = \frac{A_c^2}{4} e^{j(\Delta\hat{\omega}_c + \phi[n] - \phi[n-1])} \quad (3.26)$$

$$y[n] = \Delta\hat{\omega}_c + \phi[n] - \phi[n-1] \quad (3.27)$$

In (3.27) the phase difference is simply an approximation of the phase derivative[3], i.e., of the frequency, while $\Delta\hat{\omega_c}$ represents, if constant, a small DC component.

☐ Question 5

Save the file `toneFMmod.m` as `delayFMdemod.m`. Complete the code in the new file to implement the "Complex Delay Line Discriminator".

a) First assume that $f_0 = f_c$ and use a low-pass FIR filter with a bandwidth of about 3600 Hz. Compensate the delay introduced by the filter.

b) Plot the PSDs of the modulated signal, of $x_{bb}[n]$ and of the output. Compare in the same figure the waveforms of the modulating and demodulated signals.

c) Observe the effect on the demodulated signal of a difference of about 100 Hz between the transmitter and the receiver carriers.

■

[3] Remember the definition of a derivative of a function $\phi(t)$: $\frac{d\phi(t)}{dt} = \lim_{T_s \to 0} \frac{\phi(t) - \phi(t-T_s)}{T_s}$.

3.3 Stereo FM

Figure 3.5 illustrates the structure of the modulating signal spectrum in stereo FM as used for radio broadcasting. As it can be observed, left (L) and right (R) audio channels are not handled individually, but combined into a sum and a difference signals. This makes stereo FM trasmitters compatible with monaural receivers, since they only need to demodulate the L+R band. The L-R channel is DSB-SC modulated at a center frequency of 38 kHz and, optionally, some additional data (station name, current program, etc.) can be transmitted in the RDS band. In addition, a pilot tone is inserted at a frequency of 19 kHz, which facilitates carrier recovery and, thus, the demodulation of the L-R and RDS bands, as well as turning on the receiver stereo indicator when it is received with a sufficient power level. The value of maximum deviation frequency is typically 75 kHz. If we consider a modulating bandwidth of $W = 53$ kHz (no RDS), then, the approximate bandwidth of the FM signal, according to Carson's rule, is

$$B_{\text{FM}} = 2(\Delta f + W) = 256\,\text{kHz}. \tag{3.28}$$

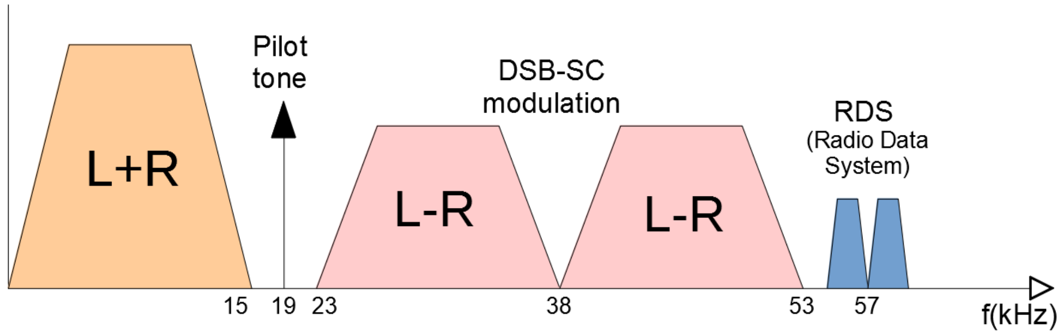

Figure 3.5: Spectrum of the modulating signal in FM stereo

☐ **Question 6**

Consider the diagram `FMReceiverSimulinkExample`. Tune in a stereo FM station (90.5 MHz, for instance) and store at least 5 seconds of the complex-baseband FM signal in the variable `xfm`.

a) Estimate the bandwidth occupied by the tuned FM station.

b) Reusing the code developed in question 5, create the script `demod_xfm` that, from `xfm`, recovers the modulating signal. Plot its spectrum.

c) Identify the different contributions in the modulating signal spectrum and estimate their bandwidth.

d) Complete the code to recover the signal that we would obtain at the output of a monaural FM receiver and listen to it to check if the process has worked.

e) Assuming that you have the L+R and L-R signals, how can you recover the L and R stereo channels?

■

Bibliography

[1] `https://es.mathworks.com/help/comm/ref/comm.fmmodulator-class.html`.

[2] A.B. Carlson, P.B. Crilly, and J.C. Rutledge. *Communication Systems: An Introduction to Signals and Noise in Electrical Communication*. McGraw-Hill series in electrical and computer engineering. McGraw-Hill, 2002.

[3] J.G. Proakis and M. Salehi. *Communication Systems Engineering*. Pearson Education. Prentice Hall, 2002.

[4] R.W. Stewart, K.W. Barlee, and D.S.W. Atkinson. *Software Defined Radio Using MATLAB & Simulink and the RTL-SDR*. Strathclyde Academic Media, 2015.