HY-530: Digital Communications

Explanatory notes on FSK modulation.

Instructor: Dr. Markos G. Troulis
mtroulis@csd.uoc.gr

Assigned on: Oct. 24 2002
Due date: Oct 24 2002

The following text has been prepared to help you to undestand from a theoretical point of view, what the FSK modulation is and to convince you about the steps to follow in order to demodulate the signal.

FSK Modulation.

In digital communication systems we always have to transmit some sequence of bits correspondig to a meaningful information message. Here we deal with FSK (Frequency Shift Keying) which is a special case of continuous phase modulation.

Generally speaking, in a phase modulation system the transmitted signal is of the form

$$s(t) = A\cos(\Phi(t;\ \mathbf{a}))$$

where A is a constant amplitude and $\Phi(t;\ \mathbf{a})$ is a continuous function of time called the instanteous phase. This function of time contains the information sequence $\mathbf{a}$, with $\mathbf{a} = (a_0,\ a_1, \ldots)$ and $a_i \in \{0,\ 1\}$. Thus $\mathbf{a}$ is the sequence of bits we wish to transmit.

In what follows we will denote by T_b the time duration of a bit. The instantaneous phase $\Phi(t;\ \mathbf{a})$ is the information bearing signal and is related to the concept of instantaneous frequency as explained below.

In FSK we define

$$\Phi(t;\ \mathbf{a}) = 2\pi \sum_k f_{a_k}\, g(t - kT_b) + \phi_0$$

where $g(t)$ is shown in Figure **??** and ϕ_0 is an arbitrary initial phase and f_{a_k} is either f_0, either f_1, according to which bit (0 or 1) is transmitted. Note also that in the interval $(kT_b, (k+1)T_b)$, the slope of the line-segments equals f_{a_k}.

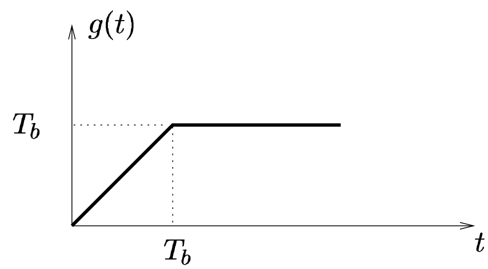

Figure 1:

Therefore $\Phi(t;\ \mathbf{a})$, looks like in Figure **??**.

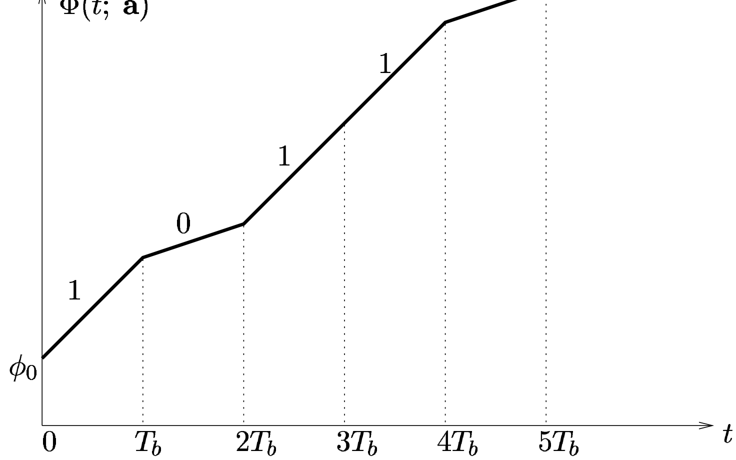

Figure 2:

In the case of FSK, when a bit 0 is transmitted, a frequency f_0 is used and for transmission of a bit 1, we use f_1. Thus for transmission of a bit a_k we use the frequency f_{a_k}.

Let's call f_c the middle frequency $f_c = \frac{f_0 + f_1}{2}$ and Δf, the frequency deviation from f_c, i.e. $\Delta f = \frac{f_1 - f_0}{2}$. Also, $f_0 = f_c - \Delta f$ and $f_1 = f_c + \Delta f$.

That way we can express f_{a_k} as:

$$f_{a_k} = f_c + b_{a_k} \Delta f$$

with $b_{a_k} = 1$ if $a_k = 1$ and $b_{a_k} = -1$ if $a_k = 0$. That way, recovering the sequence $\{b_{a_k}\}$ is equivalent to recovering the sequence $\{a_k\}$.

$\Phi(t;\ \mathbf{a})$ is now written as:

$$\Phi(t;\ \mathbf{a}) = 2\pi \sum_k (f_c + b_{a_k}\ \Delta f)\ g(t - kT_b) + \phi_0 = 2\pi f_c t + \underbrace{2\pi \sum_k b_{a_k}\ \Delta f\ g(t - kT_b)}_{\Theta(t;\ \mathbf{a})} + \phi_0$$

This shows that the signal $s(t)$ is a bandpass signal with carrier frequency equal to f_c. We will now proceed to its demodulation, recovering the in-phase and quadrature components. We have:

$$\begin{aligned} s(t) &= A\cos(2\pi f_c t + \Theta(t;\ \mathbf{a}) + \phi_0) \Rightarrow \\ s(t)\cos(2\pi f_c t) &= \frac{A}{2}\underbrace{\cos(4\pi f_c t + \Theta(t;\ \mathbf{a}) + \phi_0)}_{\substack{\text{High frequency} \\ \text{component } (2f_c)}} + \frac{A}{2}\underbrace{\cos(\Theta(t;\ \mathbf{a}) + \phi_0)}_{\substack{\text{Low frequency} \\ \text{component } (\Delta f)}} \\ s(t)\sin(2\pi f_c t) &= \frac{A}{2}\underbrace{\sin(4\pi f_c t + \Theta(t;\ \mathbf{a}) + \phi_0)}_{\substack{\text{High frequency} \\ \text{component } (2f_c)}} + \frac{A}{2}\underbrace{\sin(\Theta(t;\ \mathbf{a}) + \phi_0)}_{\substack{\text{Low frequency} \\ \text{component } (\Delta f)}} \end{aligned}$$

With a low-pass filter we eliminate the hight frequency components and then:

$$\begin{aligned} s_c(t) &= \frac{A}{2}\cos(\Theta(t;\ \mathbf{a}) + \phi_0) \quad \text{in-phase comp.} \\ s_s(t) &= \frac{A}{2}\sin(\Theta(t;\ \mathbf{a}) + \phi_0) \quad \text{quadrature comp.} \end{aligned}$$

That way:

$$\tan(\Theta(t;\ \mathbf{a}) + \phi_0) = \frac{s_s(t)}{s_c(t)} \Rightarrow \Theta(t;\ \mathbf{a}) + \phi_0 = \tan^{-1}\left(\frac{s_s(t)}{s_c(t)}\right)$$

Differentiating $\Theta(t;\ \mathbf{a})$ gives:

$$\frac{d[\Theta(t;\ \mathbf{a}) + \phi_0]}{dt} = \frac{d\Theta(t;\ \mathbf{a})}{dt} = 2\pi\Delta f \sum_k b_{a_k}\, g'(t - kT_b) = 2\pi\Delta f \sum_k b_{a_k}\, p(t - kT_b)$$

where $g'(t) = p(t)$ is a square pulse as shown in figure **??**.

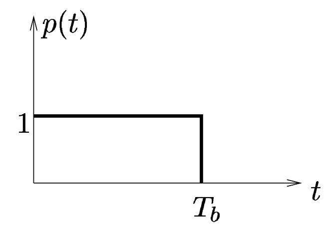

Figure 3:

That way we recover the transmitted sequence $\{a_k\}$, from the modulated pulse-train $\frac{d\Theta(t;\ \mathbf{a})}{dt}$. Schematically the situation is depicted in figures **??** and **??**:

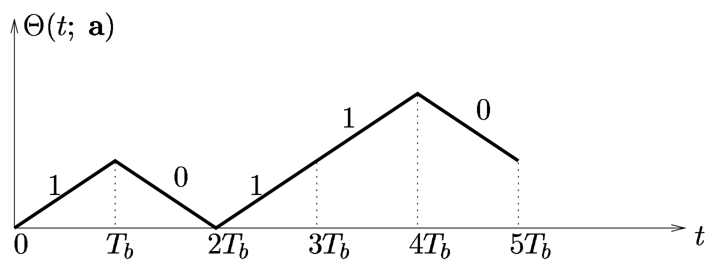

Figure 4:

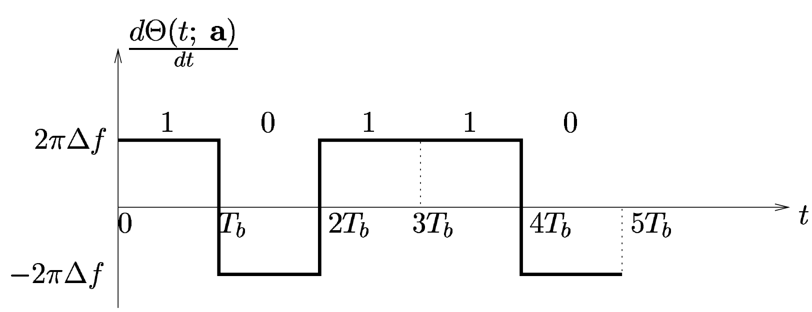

Figure 5: