Digital modulation and mobile radio (II)

2 Basic modulation techniques

The amplitude, frequency and/or phase of the RF or baseband signal described by equations (2) to (6) are modified by the data signal so that the information it represent is impressed on the RF signal. The terms M-ary amplitude, frequency or phase modulation are used to indicate which parameter is varied to represent the information. The term keying is used for digital modulation. The abbreviations ASK for amplitude-shift keying, FSK for frequency-shift keying and PSK for phase-shift keying are widely used internationally.

Straight ASK has ceased to have any significance in practice, but FSK and PSK are widely used in their binary and M-ary forms. With M-ary PSK, the upper limit for M is 8 because noise susceptibility rises disproportionally as M increases. A combination of ASK and PSK is used for M >8. This leads to M-ary quadrature amplitude modulation (QAM). In this case the carrier can assume M states or, as has already been explained, there exists a set of M time-domain signals, each of which represents a digital word ld(M) bits in length (ld = logs to base 2).

The actual modulation process involves multiplying an RF carrier of the form $A \cdot \cos[(2\pi f_c t + \varphi(t)]$, which can also be expressed in complex form as $A \cdot \exp(j2\pi f_c t) \cdot \exp[j\varphi(t)]$, by the baseband signal. A simple ring mixer can be used to perform straightforward types of modulation like binary ASK or PSK. The data sequence is converted into a unipolar (ASK) or bipolar (binary PSK) NRZ signal. With M-ary PSK and QAM, the signal to be modulated must be split into its I and Q components. Both components are modulated by two different time-domain signals $c_I(t)$ and $c_Q(t)$, which are derived from the data sequence a(n). However, it turns out that these two time-domain signals are identical to the real and imaginary parts of the complex, equivalent baseband signal to within the sign of the imaginary part. Modulation is then essentially a process of mapping the data sequence a(n) to be transmitted onto the complex envelope of the modulated RF signal.

2.1 Amplitude-shift keying

Amplitude-shift keying (ASK) is the simplest way of modulating digital information onto a carrier. Even though on/off keying (OOK), the original version of ASK, is no longer commonly used, it is the key to understanding the fundamentals of digital modulation.

To simplify the discussion, we shall assume that the data stream which modulates the (real) carrier $A \cdot \cos(2\pi f_c t)$ is a sequence a(n) with $a \in \{0, 1\}$ comprising an alternating sequence of 1s and 0s, ie a(n) = 1, 0, 1, 0, 1, 0, ... This sequence is converted into a unipolar NRZ signal by a digital interpolation filter with

$$h(t) = \begin{cases} \frac{1}{T} \text{ for } -\frac{T}{2} \le t \le \frac{T}{2} \\ 0 \text{ elsewhere} \end{cases}$$

After D/A conversion, a signal with the Fourier series

$$u(t) = \frac{1}{2} A \cdot \left[1 + \frac{2}{\pi} \cdot \cos\left(2\pi \left(\frac{1}{2} f_{bit}\right) t\right) - \frac{2}{3\pi} \cdot \cos\left(2\pi \left(\frac{3}{2} f_{bit}\right) t\right) + \dots \right] \quad (7)$$

is obtained.

The DC component is half the amplitude (A/2) and there are spectral lines at odd half-integer multiples of the bit rate, ie $(2n+1) f_{bit}/2$. There are no spectral lines at whole integer multiples of the bit rate $n \cdot f_{bit}$ $(n \neq 0)$. If the data sequence is a random, uniformly distributed sequence of 1s and 0s, a good assumption for actual data transmissions, the discrete spectrum tends to a continuous spectrum in the limit. Again there is a DC component of A/2 and zeroes at multiples of the bit rate. This signal is fed to the modulator.

It is easy to see that the carrier is, in effect, turned on and off when it is multiplied by the unipolar NRZ signal. If modulation is taken to be the assignment of a symbol a_i from an alphabet comprising N symbols to a signal s_i from a set of N signals (in this case N = 2), the signal $s_1 = \cos(2\pi f_c t)$ is assigned to the symbol 1 and the signal $s_2 = 0$ to the symbol 0.

As the carrier signal and the baseband signal are multiplied together, it follows from the shift theorem for Fourier transforms that the spectrum of the modulated RF signal is

$$s(t) = \frac{1}{2} A \cdot \left[\cos(2\pi f_c t) + \frac{1}{\pi} \cdot \cos\left(2\pi \left(f_c \pm \frac{1}{2} f_{bit}\right) t\right) - \frac{1}{3\pi} \cdot \cos\left(2\pi \left(f_c \pm \frac{3}{2} f_{bit}\right) t\right) + \dots \right] \quad (8)$$

for the sequence a(n) = 1, 0, 1, 0, 1, 0 ...... and a continuous spectrum in the form of a symmetrical mapping of the baseband spectrum into the RF band for a random sequence a(n). In both cases the carrier frequency is present in the spectrum but at half its original amplitude. The spectrum has zeroes at the frequencies $f_c \pm n f_{bit}$ (FIG 3).

Equation (8) shows that both halves of the spectrum centered on the carrier frequency would, in theory, be infinitely wide. Therefore, to ensure efficient use of the RF band, the spectrum must be bandlimited. This is best done in the baseband. Specifically, all frequencies in the baseband signal up to at least $f_{bit}/2$ must be allowed through. This is equivalent to limiting the bandwidth of the RF signal to fc ± $f_{bit}/2$, in other words to f_{bit}.

Abrupt limiting of the baseband at $f_{bit}/2$ is counterproductive because large signal delays are introduced. The filter that is used should, therefore, have a continuous transition from the passband to the stopband which follows a cosine function for example.

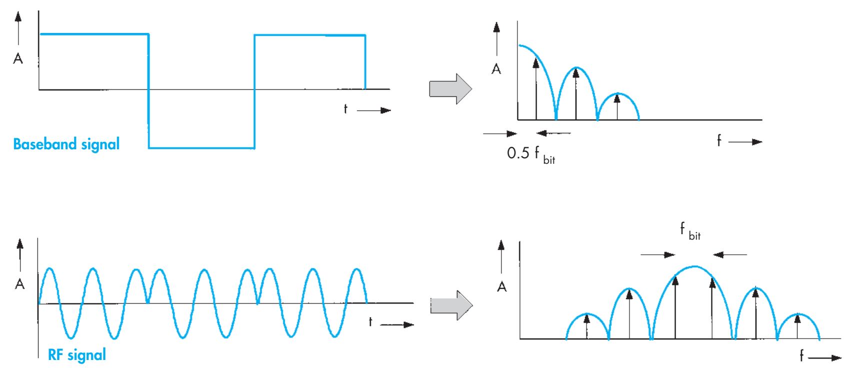

FIG 4 Time-domain signals and spectra with binary phase-shift keying

$$H(f) = \begin{cases} 1 & \text{for } 0 \le f \le \dfrac{(1-\alpha)}{2T} \\ \dfrac{1}{2}\left\{1 - \sin\left[\dfrac{\pi(2fT-1)}{2\alpha}\right]\right. & \text{for } \dfrac{(1-\alpha)}{2T} < f \le \dfrac{(1-\alpha)}{2T} \\ 0 & \text{for } \dfrac{(1-\alpha)}{2T} < f \end{cases} \qquad (9)$$

This filter replaces the sin x/x lowpass which was mentioned previously. The RF signal now has a bandwidth B which is a function of the roll-off factor α and in fact when α = 0, B = f_{bit} and when α = 1, B = $2f_{bit}$. In practice roll-off factors between 0.35 and 0.5 are widely used; this corresponds to RF bandwidths between $1.35f_{bit}$ and $1.5f_{bit}$. It is then possible to define a bandwidth efficiency which indicates what bit rate per Hz of bandwidth can be transmitted. The theoretical upper limit for ASK is 1 bit/s/Hz, but in practice efficiencies between 0.65 and 0.8 bit/s/Hz are encountered.

2.2 Binary phase-shift keying

If the data sequence a(n) is mapped to a sequence of delta functions a(n) · δ (nT) ∈ {+1;–1}, an NRZ signal is output by the sin x/x lowpass. This signal, unlike the modulation signal for ASK, is bipolar (binary phase-shift keying, BPSK). If the NRZ signal is used to modulate a carrier of the form A · $\cos(2\pi f_c t)$, an RF signal which is phase-shifted by 180° as compared to the modulating signal is obtained. This means that $s_1(t) = -s_2(t)$ and this type of modulation is called antipodal.

The (real) spectrum of the baseband signal is given by

$$u(t) = A \cdot \left[\frac{4}{\pi} \cdot \cos\left(2\pi\left(\frac{1}{2} f_{bit}\right)t\right) - \frac{4}{3\pi} \cdot \cos\left(2\pi\left(\frac{3}{2} f_{bit}\right)t\right) + \ldots\right]. \qquad (10)$$

and has no DC component.

The spectrum of the modulated carrier

$$s(t) = A \cdot \left[\frac{2}{\pi} \cdot \cos\left(2\pi\left(f_c \pm \frac{1}{2} f_{bit}\right)t\right) - \frac{2}{3\pi} \cdot \cos\left(2\pi\left(f_c \pm \frac{3}{2} f_{bit}\right)t\right) + \ldots\right], \qquad (11)$$

is obtained by multiplying the baseband signal and the carrier together. The spectrum does not contain the carrier frequency (FIG 4). The occupied bandwidth is the same as that for ASK and has been limited by passing the baseband signal through a lowpass filter with a cos roll-off before modulation. Consequently, ASK and BPSK have the same theoretical maximum bandwidth efficiency of 1 bit/s/Hz and a bandwidth efficiency between 0.65 and 0.8 bit/s/Hz in practice.

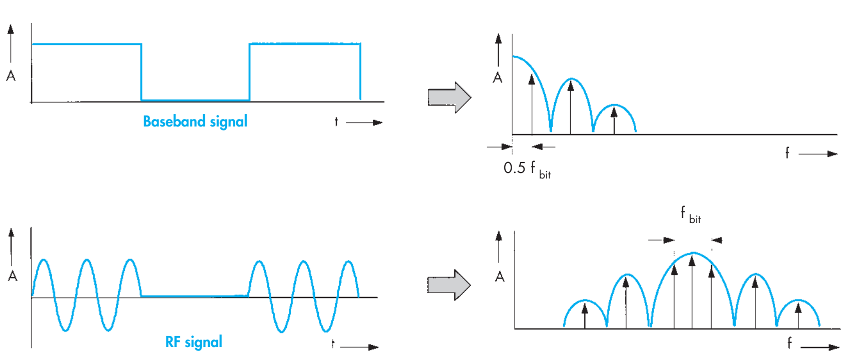

FIG 3 Time functions and spectra with amplitude-shift keying

2.3 M-ary quadrature amplitude modulation

To increase bandwidth efficiency, two, three or in general k consecutive bits from a data sequence a(n) can be combined to form a new symbol b(m) by means of serial/parallel conversion. As a result, the symbol rate is reduced to f_{bit}/k. The new symbols are referred to as dibits, tribits, quadbits or k-bit words in general. The modulation process requires $M = 2^k$ RF signals, each with a different phase and/or amplitude. The $M = 2^k$ possible symbols b(m) are mapped onto these signals.

An I/Q modulator is the best type of modulator to use. First of all, it splits the unmodulated RF signal into two components. The quadrature or Q component is phase-shifted by 90° with respect to the in-phase or I component. Therefore the unmodulated I component is described by $\cos(2\pi f_c t)$ and the unmodulated Q component by $-\sin(2\pi f_c t)$. Both components are fed to mixers, where they are multiplied with the modulating signals $c_I(t)$ and $c_Q(t)$; $c_I(t)$ and $c_Q(t)$ are derived from the symbol sequence b(m) and to within the sign of the Q component are identical to the real and imaginary parts of the complex envelope of the modulated RF signal. The products $c_I(t) \cdot \cos(2\pi f_c t)$ and $c_Q(t) \cdot [-\sin(2\pi f_c t)]$ are added together to give the modulated RF signal. The modulation process is reduced to mapping the symbol sequence b(m) onto the two baseband components. In the case of unfiltered quadrature amplitude modulation (QAM), staircase signals in the time domain with M/2 possible values are produced (FIG 5).

By passing the baseband signal through a filter with a cutoff frequency equal to half the symbol rate, the bandwidth of the RF signal is limited to the symbol rate = bit rate/k, which means that the bandwidth efficiency is k times better than for BPSK. Quadrature phase-shift keying (QPSK) is the first step in this direction. Two consecutive bits are combined to form a dibit which can represent one out of $M = 2^2 = 4$ symbols. These symbols are mapped onto the phases $\varphi_i \in \{45°, 135°, 225°, 315°\}$ of the RF signal or onto the four time-domain signals $s_i(t) = A \cdot \cos[2\pi f_c t + (2i + 1) \cdot \pi/4]$ with $i \in \{0, 1, 2, 3\}$. TABLE 1 shows how the bit sequences, the dibits, the modulation signals and the phases of the RF signal are related. FIG 6 shows the signals in the time domain, in the frequency domain and by means of a phase state diagram. It should be noted that in this case too there is no carrier-frequency component in the spectrum.

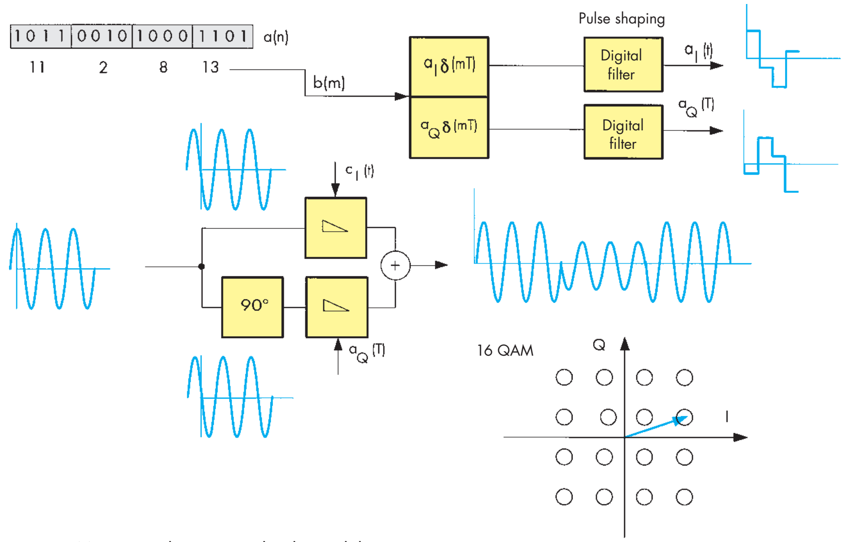

FIG 5 M-ary quadrature amplitude modulation

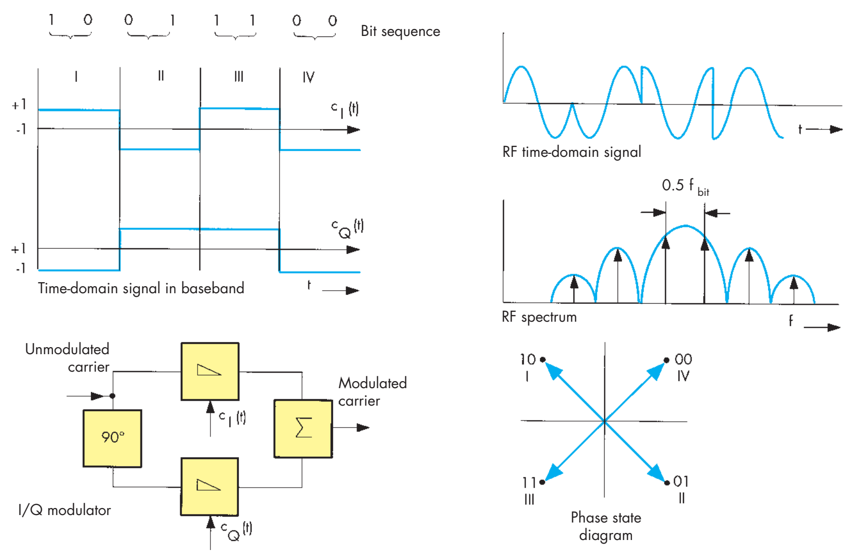

FIG 6 Quadrature phase-shift keying

Bit sequence	Dibit	$c_I(t)$	$c_Q(t)$	φ
00	I	1	1	45°
10	II	–1	1	135°
11	III	–1	–1	225°
01	IV	1	–1	315°

TABLE 1 Modulation parameter assignments

Theoretically, the bandwidth efficiency can be increased to 2 bit/s/Hz and so is twice the maximum bandwidth efficiency for ASK and BPSK. In practice, values between 1 and 1.5 bit/s/Hz can be obtained.

To be continued Peter Hatzold