# INTRODUCTION TO DIGITAL MODULATION SCHEMES

**Geoff Smithson.**

## Overview

The choice of digital modulation scheme will significantly affect the characteristics, performance and resulting physical realisation of a communication system. There is no universal 'best' choice of scheme, but depending on the physical characteristics of the channel, required levels of performance and target hardware trade-offs, some will prove a better fit than others. Consideration must be given to the required data rate, acceptable level of latency, available bandwidth, anticipated link budget and target hardware cost, size and current consumption. The physical characteristics of the channel, be it hardwired without the associated problems of fading, or a mobile communications system to a F1 racing car with fast changing multipath, will typically significantly affect the choice of optimum system.

The objective of this paper is to review the key characteristics and salient features of the main digital modulation schemes used, including consideration of the receiver and transmitter requirements. Simulation is used to compare the performance and tradeoffs of three popular systems, MSK, GMSK and BPSK, including analysis of key parameters such as occupied bandwidth and Bit Error Rate (BER) in the presence of Additive White Gaussian Noise (AWGN). Finally the digital modulation schemes used in the current and proposed cellular standards are summarised.

## Digital Modulation

Firstly, what do we mean by digital modulation? Typically the objective of a digital communication system is to transport digital data between two or more nodes. In radio communications this is usually achieved by adjusting a physical characteristic of a sinusoidal carrier, either the frequency, phase, amplitude or a combination thereof. This is performed in real systems with a modulator at the transmitting end to impose the physical change to the carrier and a demodulator at the receiving end to detect the resultant modulation on reception.

## Noise

All real systems contain noise. This noise will result in a non zero probability of an error occurring at the demodulator. Again as different systems have different sensitivities to errors different ways of handling them exist. An error could potentially have a much more significant effect in a system that remotely controls of a piece of machinery, for example a crane, than in a consumer voice system. Other parameters such as the propagation channel and acceptable data latency will affect the optimum choice for the handling of errors.

## Propagation Characteristics

Radio communication systems come in a variety of different forms and consequently have very different channel characteristics. Microwave point to point links are typically designed with slim margins as the channel will vary little over time, with the exception of possibly increased loss from rain. The dynamic range required at the receiver will consequently be small.

Geoff Smithson is an RF Engineer with Plextek Ltd, London Road, Great Chesterford, Essex.
Tel: +44 (0)1799 533200. Fax: +44 (0)1799 533201. Email: gs@plextek.co.uk

However, mobile communication systems operating in the VHF and UHF bands are typically not line-of-sight (LOS) and therefore consist of the sum of signals from different paths. Potentially this will result in intersymbol interference due to delay spreads as well as considerable amplitude variations with time and/or distance. Physical objects will also cause 'shadowing' resulting in loss of signal, for instance a vehicle entering a tunnel or traveling behind a building. Additionally, due to the varying proximity between two nodes, a considerably larger dynamic range will be required at the receiver together with an increased link budget margin.

It is clear that these two communications systems will require different properties from a modulation scheme.

**Error Handling**

The existence of Noise as well as the variations in propagation characteristics of some communication channels makes the correct handling of errors vital. Some of the ways in which errors are handled include, error detection, error correction, data acknowledgment and data re-sends. Techniques for error correction, such as performing a Cyclic Redundancy Check (CRC) on the data can guarantee detection of errors with vastly greater certainty than a typical system BER. This detection process can then be used, if required, to send an acknowledgment (ACK) or a failed acknowledgment (NAK) back to the transmitting node to request re-transmission. Typically a system based on a combination of error detection and acknowledgments can perform well in a stationary environment where data latency is not a problem. The principle advantages of this type of system are the simplicity and the high data through-put relative to the data rate for good SNR conditions. This type of system doesn’t, however, tend to work well in poor SNR conditions due to the large number of re-sends necessary. In order to improve the typical BER of a mobile communications systems error detection and correction is usually employed. This involves the transmission of some additional data in order to allow the detection, identification and correction of errors.

**<u>Modulation Schemes</u>**

**Coherent and Incoherent Systems.**

The terms coherent and incoherent are frequently used when discussing the generation and reception of digital modulation. When linked to the process of modulation the term coherence relates to the ability of the modulator to control the phase of the signal, not just the frequency. For example Frequency Shift Keying (FSK) can be generated both coherently with an IQ modulator and incoherently with simply a Voltage Controlled Oscillator (VCO) and a digital voltage source, as shown in figure 1. below.

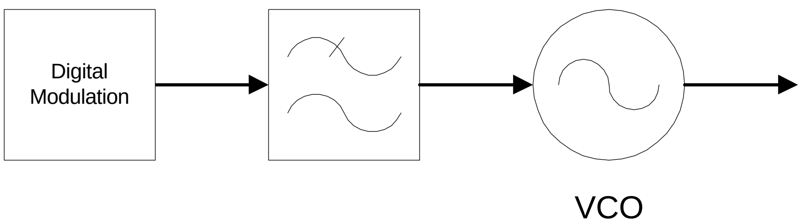


**Figure 1(a). In-coherent generation of FSK.**

With the system in figure 1(a) the instantaneous frequency of the output waveform is determined by the modulator (within a tolerance set by the VCO and data amplitude etc) but the instantaneous phase of the signal is not controlled and can have any value. Alternatively coherent generation of modulation is achieved as shown in figure 1(b). Here the phase of the signal is controlled, rather than the frequency.

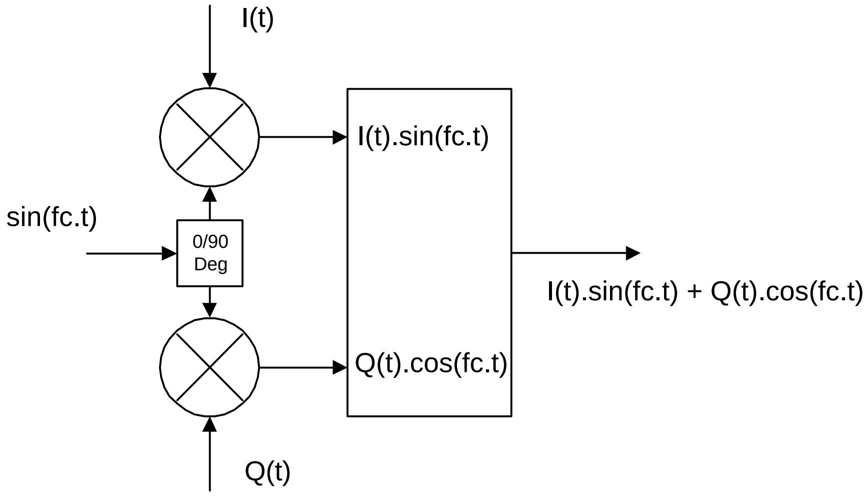


**Figure 1(b). Coherent Generation of FSK**

When a coherent modulator is used to generate FSK the exact signal frequency and phase are controlled. The modulator shown above offers the possibility to shape the resultant carrier phase trajectory at baseband either with analogue filtering or digital signal processing and a DAC. This can be used to generate both constant amplitude and amplitude modulated signals.

Use of the term coherent with respect to the act of demodulation refers to a system that makes a demodulation decision based on the received signal phase, not frequency. The additional 'information' available results in an improved BER performance. The high level of digital integration now possible in semiconductor devices has made digitally based coherent demodulators common in mobile communications systems.

**Frequency Shift Keyed (FSK)**

As previously stated applying modulation in wireless communications involves modifying the phase or amplitude, or both, of a sinusoidal carrier. One of the simplest, and widest used system, is frequency modulation. This exists in a great variety of forms, as will be discussed later, but in essence involves making a change to the frequency of the carrier to represent a different level. The generic name for this family of modulation is Frequency Shift Keying (FSK).

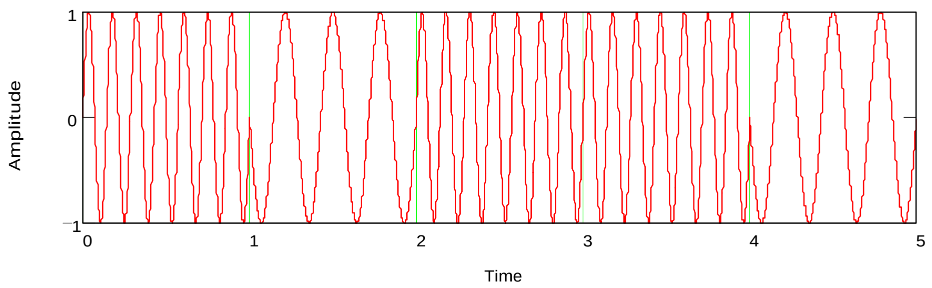


**Figure 2. Binary (2 level) FSK modulation.**

FSK has the advantage of being very simple to generate, simple to demodulate and due to the constant amplitude can utilise a non-linear PA. Significant disadvantages, however, are the poor spectral efficiency and BER performance. This precludes its use in this basic form from cellular and even cordless systems.

**Minimum Shift Keyed (MSK)**

Minimum Shift Keying is FSK with a modulation index of 0.5. Therefore the carrier phase of an MSK signal will be advanced or retarded 90° over the course of each bit period to represent either a one or a zero. Due to this exact phase relationship MSK can be considered as either phase or frequency modulation. The result of this exact phase relationship is that MSK can't practically be generated with a voltage controlled oscillator and a digital waveform. Instead an IQ modulation technique, as for PSK, is usually implemented.

Coherent demodulation is usually employed for MSK due to the superior BER performance. This is practically achievable, and widely used in real systems, due to the exact phase relationship between each bit. The spectral characteristics and BER performance of MSK are considered later.

**Gaussian Minimum Shift Keyed (GMSK)**

A variant of MSK that is employed by some cellular systems (including GSM) is Gaussian Minimum Shift Keying. Again GMSK can be viewed as either frequency or phase modulation. The phase of the carrier is advanced or retarded up to 90° over the course of a bit period depending on the data pattern, although the rate of change of phase is limited with a Gaussian response. The net result of this is that depending on the Bandwidth Time product (BT), effectively the severity of the shaping, the achieved phase change over the bit may fall short of 90°. This will obviously have an impact on the BER, although the advantage of this scheme is the improved bandwidth efficiency. The extent of this shaping can clearly be seen from the 'eye' diagrams in Figure 3 below for BT=0.3, BT=0.5 and BT=1.

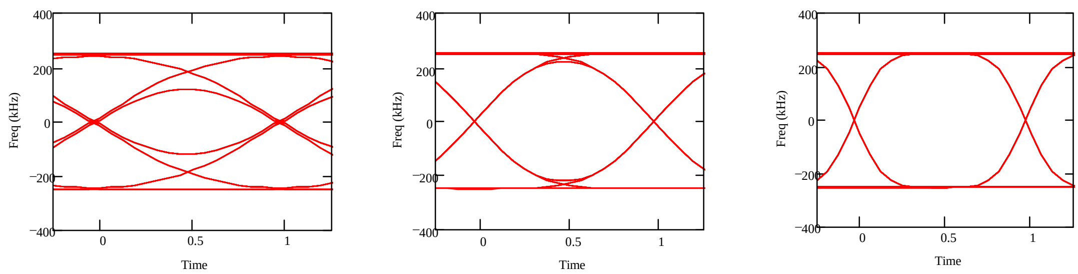


**Figure 3. Eye Diagrams for GMSK with BT=0.3 (left), BT=0.5 (centre) and BT=1 (right).**

This resultant reduction in the phase change of the carrier for the shaped symbols (ie 101 and 010) will ultimately degrade the BER performance as less phase has been accrued or retarded therefore less noise will be required to transform a zero to a one and vice versa. The principle advantages of GMSK, however, are the improved spectral efficiency and constant amplitude. The resulting signal spectras for BT= 0.3, 0.5, 1 and MSK are shown below in Figure 4.

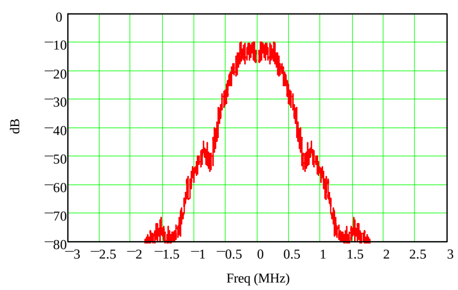


**Figure 4(a). BT=0.3**

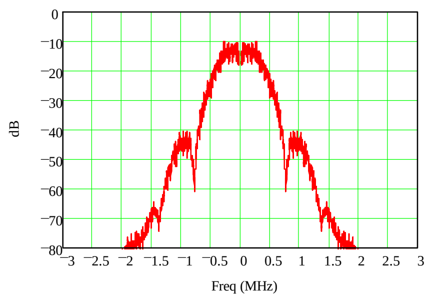


**Figure 4(b). BT=0.5**

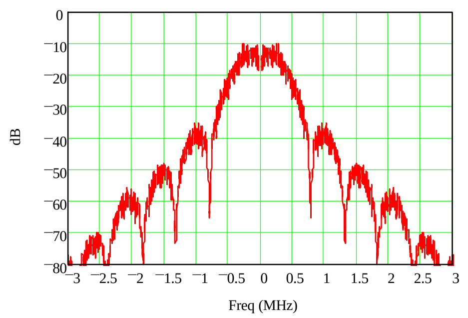


**Figure 4(c). BT=1**

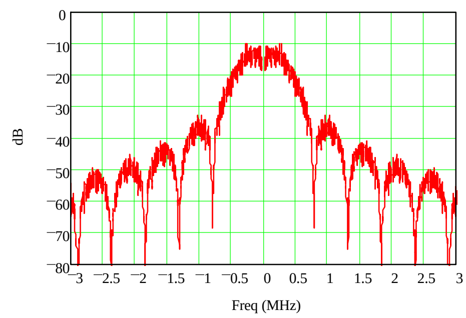


**Figure 4(d). MSK**

All the waveforms displayed above (GMSK and MSK) have constant amplitude. That is to say that their quadrature phase trajectory never leaves the unit circle. This can be a significant property, particularly as it allows the Power Amplifier device to be operated further into compression yielding improved efficiency and increased output power, without significant spectral re-growth.

**Phase Shift Keyed (PSK)**

An alternative to imposing the modulation onto the carrier by varying the instantaneous frequency is to modulate the phase. This can be achieved simply by defining a relative phase shift from the carrier, usually equi-distant for each required state. Therefore a two level phase modulated system, such as Binary Phase Shift Keying, has two relative phase shifts from the carrier, + or - $90^o$. Typically this technique will lead to an improved BER performance compared to MSK. The resulting signal will, however, probably not be constant amplitude and not be very spectrally efficient due to the rapid phase discontinuities. Some additional filtering will be required to limit the spectral occupancy. Phase modulation requires coherent generation and as such if an IQ modulation technique is employed this filtering can be performed at baseband.

**Binary Phase Shift Keyed (BPSK)**

The simplest form of phase modulation is binary (two level) phase modulation. With theoretical BPSK the carrier phase has only two states, +/- $\pi/2$. Obviously the transition from a one to a zero, or vice versa, will result in the modulated signal crossing the origin of the constellation diagram resulting in 100% AM. Figure 5(a) below shows the theoretical spectra of a 1 Mbits BPSK signal with no additional filtering. Several techniques are employed in real systems to improve the spectral efficiency. One such method is to employ Raised Cosine filtering. Figure 4(b) below shows the improved spectral efficiency achieved by applying a raised cosine filter with $\beta$=0.5 to the baseband modulating signals.

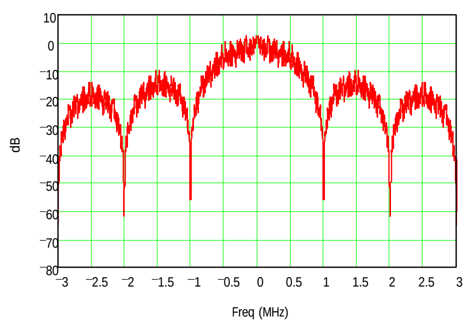


**Figure 5(a). Theoretical BPSK**

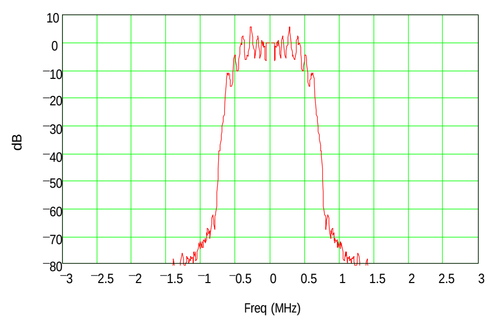


**Figure 5(b). Raised Cosine BPSK $\beta$=0.5**

The improved spectral efficiency will result in some closure of the eye as can be seen in figure 6a and 6b.


**Figure 6(a). Theoretical BPSK** **Figure 6(b). Raised Cosine BPSK β=0.5**

One potentially undesirable feature of BPSK that the application of a raised cosine filter will not improve is the 100% AM. In a real system the shaped signal will still require a linear PA to avoid spectral re-growth. Further hybrid versions of BPSK are used in real systems that combine constant amplitude modulation with phase modulation. One such example would be Constant Amplitude '50%' BPSK, generated with shaped I and Q vectors designed to rotate the phase around the unit circle between the two constellation points. For a 010 data sequence the trajectory spends 25% of the time travelling from one point to other, 50% of the time at the required point and 25% of the time returning. The resulting carrier phase shift is shown in Figure 7 below.

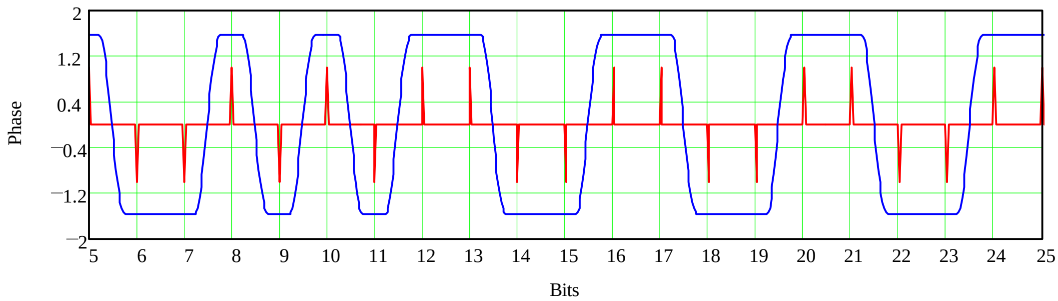


**Figure 7. Constant Amplitude '50%' BPSK.**

**Quadrature Phase Shift Keyed (QPSK)**

Higher order modulation schemes, such as QPSK, are often used in preference to BPSK when improved spectral efficiency is required. QPSK utilises four constellation points, as shown in figure 8 below, each representing two bits of data. Again as with BPSK the use of trajectory shaping (raised cosine, root raised cosine etc) will yield an improved spectral efficiency, although one of the principle disadvantages of QPSK, as with BPSK, is the potential to cross the origin, hence generating 100% AM.

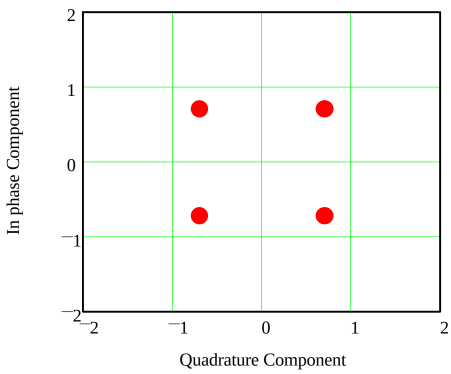


**Figure 8. Constellation points for QPSK.**

**Π/4-Quadrature Phase Shift Keyed (π/4-QPSK)**

A variant of QPSK that is employed in several digital systems is π/4-QPSK. As with QPSK two bits are coded onto each symbol, although the quadrature constellations for adjacent bits are offset by π/4 radians. The two sets of constellation points are shown in figure 9. below.

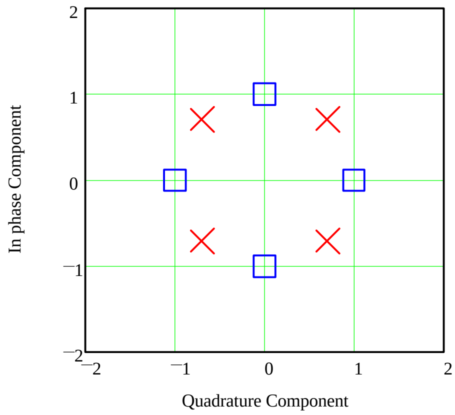


**Figure 9. Constellation points for π/4-QPSK.**

One advantage of π/4-QPSK is the improved spectral efficiency, compared to MSK and GMSK, particularly when used with raised cosine phase trajectory shaping due to coding two bits per symbol. Additionally the phase trajectory will no longer cross the origin, avoiding the generation of 100% AM, allowing a harder saturation mode of operation for the PA. Published data suggests that the BER performance will be slightly degraded from that of QPSK. (Approximately 1 dB @ $1x10^{-4}$ BER [2]).

**Offset Quadrature Phase Shift Keyed (O-QPSK)**

The final variant of QPSK to be considered is Offset Quadrature Shift Keying, or O-QPSK. As previously discussed the potential for a $180^o$ phase shift in QPSK results in the requirement for better linearity in the PA and the potential for spectral re-growth due to the 100% AM. O-QPSK reduces this tendency by adding a time delay of one bit period (half a symbol) in the Q arm of the modulator. The result is that the phase of the carrier is potentially modulated every bit (depending on the data), not every other bit as for QPSK, hence the phase trajectory never approaches the origin. The ability of the modulated signal to demonstrate a phase shift of $180^o$ is therefore removed.

As with the other phase modulation schemes considered, shaping of the phase trajectory between constellation points is typically implemented with a raised cosine filter to improve the spectral efficiency. Due to the similarities between QPSK and O-QPSK similar signal spectra and probability of error are achieved. O-QPSK is utilised in the North American IS-95 CDMA cellular system for the link from the mobile to the basestation.

**Communications Link BER Simulations**

The BER results in the following sections have been produced by performing computer based simulation of the modulator, channel (AWGN), as well as bandpass filtering and demodulation. The assumption of perfect carrier and bit synchronisation has been made. This is a very significant assumption and is unacceptable for a real system design, particularly at low to moderate BER. Ideally the addition of carrier and bit synchronisation algorithms should be implemented in the simulator. Obviously the performance of these will vary with physical realisation. A standard approximation for Gaussian noise has been used to convert the uniformly distributed random variables generated by the computer to Gaussian random variables.

In a real mobile communications system the link between a moving (particularly vehicular) node and a basestation will be subject to Rayleigh fading, resulting in fast phase shifts. This will have a significant affect on the resultant BER performance, dependant on demodulator implementation and channel, possibly increasing the required C/N for a specific BER by as much as 10 - 15 dB.

Considerable caution must be taken when comparing graphs of BER v Eb/No. As with most such graphs the ones presented in figure 10 below are calculated using perfect carrier and bit synchronisation. This is not realistic, particularly under low to medium BER conditions. Additionally the simulations have been performed with a relatively wideband input to the demodulator. Due to the varying spectral occupancy of different modulation schemes it maybe possible to employ narrower channel filtering for some schemes rather than others. For example GMSK with a BT=0.3 occupies less spectrum than MSK.

**MSK** The simulated BER performance for MSK is displayed in figure 10 (a) below.

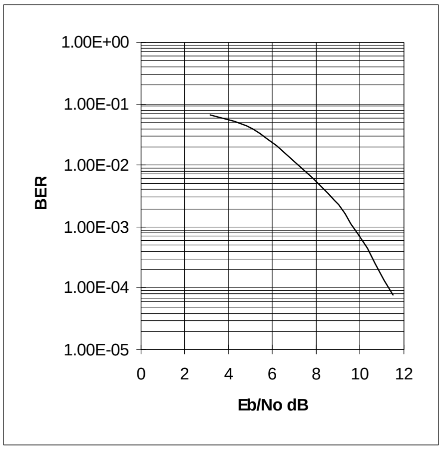


**Figure 10(a). Simulated Bit Error Rate for MSK.**

**GMSK** The simulated BER performance for GMSK is displayed in figure 10 (b) below.

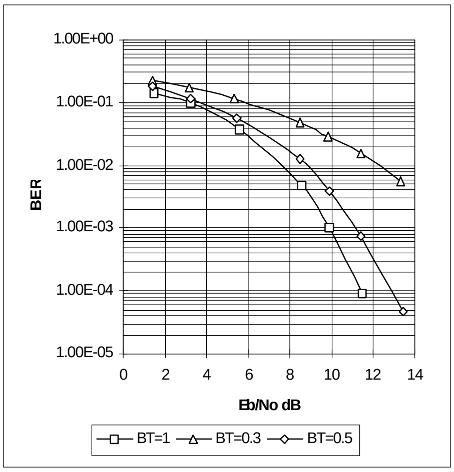


**Figure 10(b). Simulated Bit Error Rate for GMSK.**

The simulated performance of GMSK with a BT=0.3, shown above, is considerably poorer than the C/N and BER numbers achieved for implementations of this scheme in real systems (for example GSM). This is an indication of the considerable improvement to the BER performance that can be achieved with techniques such as the implementation of Maximum Likelihood algorithms [1] and optimum DSP based demodulators and filters.

**BPSK**

The simulated BER performance for BPSK is displayed in figure 10(c). below.

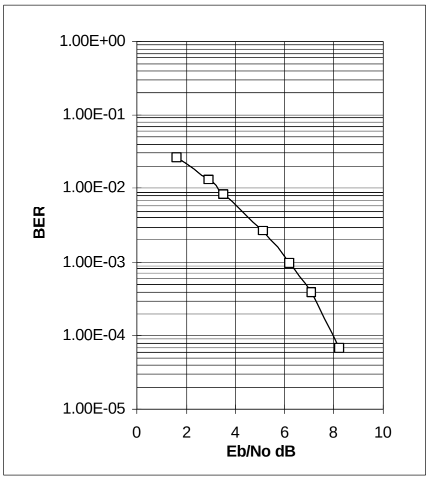


**Figure 10(c). Simulated Bit Error Rate for BPSK**

**Summary of Current and Proposed Digital Mobile and Cellular Systems**

| **Standard:** | **Principle region of operation:** | **Access method:** | **Modulation Scheme:** | **System Characteristics** |
|---|---|---|---|---|
| | | | | |
| CT2 | World wide | FDMA | GFSK (BT=0.3) | Poor range |
| DECT | Europe | TDMA | GFSK (BT=0.5) | High bit rate. Only suitable for fixed channel |
| PHS | Japan | TDMA | D $\pi$/4 QPSK ($\alpha$=0.5) | Long battery life supporting very small handsets due to system & protocol design |
| IS-54 | USA | TDMA | $\pi$/4 QPSK ($\alpha$=0.35) | |
| GSM | Worldwide | TDMA | GMSK (BT=0.3) | Widely used system |
| IS-95 | USA + | CDMA | QPSK & OQPSK | |
| PDC | Japan | TDMA | $\pi$/4 QPSK ($\alpha$=0.5) | |
| UMTS | Europe + | CDMA\TDMA | TBD* | |
| IMT-2000 | Worldwide | CDMA\TDMA | TBD* | |

*TBD To be determined.

**References**


[1] Feher; Applications of Digital Wireless Technologies to Global Wireless Communications; Prentice Hill1997; ISBN 0-13-214272-4.

[2] Lawrence E Larson; RF and Microwave Circuit Design for Wireless Communications; Artech House; 1996; ISBN 0-89006-818-6.