CHAPTER 6

PHASE SHIFT KEYING (PSK)

\[ y(t) \]
6.1 INTRODUCTION

Phase shift keying (PSK) is a very efficient digital modulation method, widely used in modern digital communication systems, such as satellite links, wideband microwave radio relay systems, etc.

In PSK, digital information is encoded in the phase function of a constant amplitude carrier signal. The information can be carried by the absolute phase difference between the modulated carrier and an unmodulated reference, or by changes in the carrier phase; the latter method is called “differential encoding”, because what matters is the phase difference between successive symbol intervals. For example, in binary PSK, “0” can be encoded as 0° phase, “1” – as 180°. Differentially encoded binary PSK will encode “0” as no change in phase, “1” – as a 180° phase change.

It is possible to have multi-level PSK: for example, the widely used quadrature (four-level) PSK utilizes phase shifts of 0°, 90°, 180°, and 270°, whereas 16-level PSK utilizes phase shifts of \[ k \times 22.5° \], where \( k = 0, 1, \ldots, 15 \).

As explained in Chapter 4, a typical set of signals for PSK is given by:

\[
S_i(t) = \begin{cases} 
\sqrt{2E/T} \cos \left( \omega_0 t + \frac{2\pi i}{m} \right) & 0 \leq t \leq T \\
0, & \text{elsewhere}
\end{cases} \tag{6-1}
\]

Where:

- \( i = 1, 2, \ldots, m \); \( m \) – the number of different signals contained in the signal set,
- \( E \) = energy of signal \( S_i(t) \),
- \( T \) = symbol duration,
- \( \omega_0 = \frac{2\pi n_0}{T} \), where \( n_0 \) is some fixed integer.

Using well-known trigonometric identities, a phase-shifted sinusoid can be broken down into a sinusoid and a cosinusoid. This enables us to find the unit functions for the PSK signal space:

\[
f_1(t) = \sqrt{2/T} \cdot \cos \omega_0 t \tag{6-2a}
\]

\[
f_2(t) = \sqrt{2/T} \cdot \sin \omega_0 t \tag{6-2b}
\]

The signal space of PSK is therefore two-dimensional (a plane), except that the signal space of binary PSK is uni-dimensional, because binary PSK has only two signals, 180 degrees apart. Signal spaces for binary PSK and for Q PSK (quadrature PSK) are shown in Fig. 4.6 and 4.7, respectively. The signal space does not change when differential encoding is used, because differential encoding is only a rule for selecting the sequence of signal points used to represent a certain sequence of data symbols.
6.2 DETECTION OF PSK

There are two practical methods for PSK detection:

- Coherent detection, where a reference carrier signal held by the receiver is compared with the received modulated signal; any phase changes are then converted into proportional voltage changes, which can be detected by a PAM detector (see Chapter 2).
- Incoherent detection, where the difference in phase between successive symbol intervals is evaluated and converted into proportional voltage changes. This mode of detection is often called differentially coherent detection. In this case differential encoding is required at the transmitter, otherwise it would be very difficult, from a practical point of view, to decode correctly the received signal.

Coherent detection of PSK signals results in the best performance, and as such, the error rate vs signal/noise ratio of a coherent PSK detector is widely used as a comparison standard for other types of modulation and demodulation methods. However, coherent detection requires very complex circuits, and in many applications, it is replaced by the less efficient incoherent detection.

6.2.1 COHERENT DETECTION

Coherent detection requires that a replica of the original transmitted carrier signal, termed the reference signal, which must be perfectly aligned in frequency and phase with the received signal, be available at the receiver. The common means for generating the reference signal is by a phase-lock loop—the carrier recovery PLL. The performance of a practical PSK system depends largely on the performance of this PLL. In the following discussion, we will first assume that a perfect reference signal is indeed available to the receiver.

When a perfect detector and reference signal are available, the error rate of a coherent detection PSK can be calculated, based on the observation that an error will be made only when the noise changes the phase of the original signal sufficiently to bring it near to another signal point. This can be understood by referring to Fig. 4.7 showing the signal space of four-level (quadrature) PSK. Four signal points exist, spaced 90° apart in a circle whose radius is proportional to the received signal strength. Noise has the effect of changing the instantaneous phase of the received signal, however in practical cases the change is normally rather small. Thus, we expect that the actual signal phase will deviate from its optimum position. Under these conditions, we must divide the signal plane into sections, each section containing those points which are nearest to one of the original signal points. We need not concern ourselves with changes in amplitude, because in a practical system these changes are removed by a limiter.

From geometrical considerations, the best division appears to be that made by bisecting the angles between adjacent signal points; this division is also marked on Fig. 4.7. The detection rule can now be simply stated:

"Measure the angle between the received signal and each signal point, and select the nearest signal point."

The reason behind this detection rule is that small noise excursions are by far more probable, therefore, it is very likely that noise has not moved the observed signal point beyond the boundaries of the section of the plane defined above.
When it is assumed that an error is made only when the observed signal point has moved outside its section, the error rate for coherently detected binary PSK is given by:

Binary PSK: \[ p_e = \frac{1}{2} \text{erfc} \sqrt{2S/N} \] (6-3)

The error rate for QPSK is given by:

Quadrature PSK: \[ p_e = \text{erfc} \sqrt{S/N} - \frac{1}{4} \text{erfc}^2 \sqrt{S/N} \] (6-4)

Where:

\( S/N \) = signal-to-noise ratio.

Fig. 6.1 shows the error rate — vs signal-to-noise ratio for various coherent PSK systems with \( m = 2, 4, 8, \) and \( 16 \) (\( m \) — number of levels).

![Figure 6.1: Symbol error rates for coherent PSK](image)

A useful approximation of the error rate of a binary coherent PSK system operating at high signal/noise ratios is given by:

Binary PSK: \[ p_e \approx e^{-S/N} \] (6-5)

A similar approximation is available for multi-level systems:

Multilevel PSK: \[ p_e \approx e^{-(S/N) \sin^2 (\pi/m)} \] (6-6)

Both equations (6-5) and (6-6) yield reasonably accurate results at S/N ratios of 10dB or higher, therefore they can be used in most practical cases.
6.2.2 DIFFERENTIALY COHERENT DETECTION OF PSK

In differentially coherent detection, the phase of the signals received during successive symbol intervals is compared, and the direction and magnitude of the phase change is evaluated. Since both compared signals are contaminated by noise, it appears that, under similar conditions, differentially coherent PSK (DCPSK) would suffer a higher error rate. In fact, since the noise powers in the two signals tend to add during the phase comparison process, it is expected that DCPSK will require a signal/noise ratio twice as good (3dB higher) than coherent PSK, in order to achieve the same error rate. Exact calculation of the error rate of DCPSK involves expressions which have no closed-form, therefore they can only be numerically evaluated by computers. However, there are several useful closed-form approximations; one of these approximations, which is suitable for calculation of error rate at medium and high signal/noise ratios, is given below:

Binary DCPSK:

$$P_e = e^{-2S/N}$$  \(\text{(6-7)}\)

Equation (6-7) can be adapted to multi-level DCPSK:

Multi-level DCPSK:

$$P_e = e^{-2(S/N) \sin^2 (\pi/2m)}$$  \(\text{(6-8)}\)

Where:

\(m = \text{number of levels; } m = 2, 3, 4, \ldots\)

The resulting curves of error rate vs signal/noise ratio are shown in Fig. 6.2.

![Figure 6.2: Symbol error rate for differentially and coherently detected PSK](image)

Comparing equation (6-6) with equation (6-8), it is seen that the difference between expressions for coherent and incoherent detection is in the exponent; the ratio of the exponents \(r\) is given in equation (6-9) below.

184
The value of $r$, in dB gives the degradation caused by incoherent detection relative to coherent detection at medium and high signal/noise ratio; Fig. 6.3 shows the dependence of $r$ on the number of levels.

Analysis of Fig. 6.3 indicates that when the number of levels is large, the degradation approaches 3dB. In particular, the degradation suffered by the widely used quadrature PSK is 2.3 dB. However, it seems rather surprising that for the binary case ($m = 2$), no degradation is indicated, at least not when high signal to noise ratios are assumed. A more exact calculation, based on numerical evaluation of the exact rate expressions, indicates that the degradation varies from 1.8 dB, at an error rate of $10^{-2}$ ($S/N = 6.1$ dB vs 4.2 dB) to 1 dB at an error rate of $10^{-4}$ (9.2 dB vs 8.2 dB), and to 0.8 dB at $10^{-6}$ (11.2 dB vs 10.4 dB).

6.3 SPECTRUM OF PSK

The spectrum of PSK is given by:

$$W(f) = T \left( \frac{\sin \pi f T}{\pi f T} \right)^2 = \frac{1}{R} \left( \frac{\sin \pi f / R}{\pi f / R} \right)^2$$

(6-10)

Where:

$T = \text{symbol interval,}$

$R = \frac{1}{T} = \text{symbol rate.}$

For binary PSK, the bit and symbol rates are equal; for quadrature PSK, the symbol rate is half the bit rate, for eight-level PSK, the symbol rate is one-third the bit rate, and so on. One interesting observation shows that using QPSK instead of binary PSK, enables transmission at twice the bit rate of BPSK in the same frequency band. Fig. 6.4 shows the normalized spectrum of PSK.
The spectrum consists of a main lobe, whose width equals twice the symbol rate, and of secondary side lobes, symmetrically positioned around the main lobes; the lobes are separated by zeros, at multiples of the transmission rate. The rate of decrease in the amplitude of the side lobes is very slow (proportional to $1/f^2$), e.g. the nearest lobe is attenuated only 13.5 dB relative to the main lobe. The slow reduction in the side lobe level is emphasized in Fig. 6.5, which shows the spectrum on a logarithmic scale.

From a practical point-of-view, the slow decrease in the side lobe level means that receivers operating on nearby frequencies will be disturbed; therefore, it becomes necessary to filter the output signal of PSK transmitters. The effects of filtering the signal is analyzed in the following paragraphs, however.
one effect needs no further explanation. Removing side lobes means that part of the transmitter power (the part contained in the side lobes) is lost. Therefore, system performance suffers, i.e. more transmitter power is required to achieve a given signal/noise ratio at the receiver antenna.

6.4 IMPLEMENTATION OF PSK TRANSMITTER

The PSK transmitter consists of three principal blocks:

1. Modulator — receives the digital data stream and a carrier signal, and generates a modulated signal whose phase changes according to the coding rule adopted.
2. Power amplifier — amplifies the modulated signal to the required level.
3. Transmitter filter — reduces harmonics and side lobes to the specified levels.

The implementation of each block will be discussed below.

6.4.1 PSK MODULATOR

6.4.1.1 Binary PSK Modulator
In principle, the basic binary PSK modulator consists of a switch, which selects the carrier or the inverted carrier, for transmission; the switch is controlled by the data stream. The function is shown in block diagram form in Fig. 6.6.

```
\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{binary_psk_modulator.png}
\caption{Block diagram for binary PSK modulator}
\end{figure}
```

The actual implementation of a binary PSK modulator depends on a number of factors, amongst them, operating frequency and data rate, and cost of implementation, are generally the dominant factors. For moderate data rates and frequencies not higher than several tens of megahertz, a transistor switch or CMOS transmission gate can provide very good performance at low cost. At higher frequencies and/or data rates, double-balanced diode mixers are generally used.
6.4.1.2 Differential Encoding
In most applications, differential encoding is required. To obtain differentially encoded binary PSK, an appropriate encoder must be inserted in the data path, before the modulator input. The logic diagram of a differential encoder is shown in Fig. 6.7.

(a) Differential encoder

From the typical waveforms given in Fig. 6.7b, it can be seen that the encoder converts the input data stream to a different data stream, so that "1"'s in the input data are converted to 180° changes in the phase of the signal generated by the modulators shown in Fig. 6.6; "0"'s in the input data leave the modulator's output phase unchanged. Since now data bits are represented by phase changes or lack thereof, it is no longer necessary to have at the receiver a replica of the transmitter carrier with which to compare the phase of the received signal. This is a necessary condition for differentially coherent detection; however, even when coherent detection is used, it is often very difficult to regenerate the carrier with the exact phase relationship required for unambiguous data demodulation (see paragraph dealing with carrier recovery at the receiver); this deficiency can be corrected by sending differentially encoded data even when coherent detection is used. Today, differential coding is universally used on PSK equipment.

6.4.1.3 Multi-level PSK Modulators
The binary modulator operates by selecting under control of the input data stream, the appropriate carrier phase. This principle can be extended to implement m-level PSK modulators, as shown in Fig. 6.8.
The modulator consists of a bank of switches, which connect the desired phase-shifted carrier to the output line, under control of the decoder circuit. The task of the decoder is to analyze the input data and determine the output signal phase to be transmitted, according to the selected encoding rule. The general implementation of Fig. 6.8 can handle any integer number of levels, 'm', but, in practice, most PSK systems are designed with \( m = 2^k \), where \( k = 2, 3, 4, \ldots \). In these particular cases, it is customary to build the multi-level PSK modulator as a bank of binary PSK modulators, fed by a network of phase shifters. Examples of quadrature (four-level) and eight-level PSK modulators are shown in Fig. 6.9, 6.10, and 6.12 respectively.

The QPSK modulator shown in Fig. 6.9 transmits two bits at a time. A pair of data bits is temporarily stored in a register; one of the bits is applied to the in-phase channel binary modulator, the other bit – to the quadrature channel modulator. The designations “in-phase” and “quadrature” refer to the phase of the carrier signal fed to the modulators, relative to the phase of the unmodulated carrier, as can be seen in Fig. 6.9.
FIGURE 6.10: Alternate implementation for QPSK modulator

Each modulator channel receives, therefore, a binary data stream, at half the input data rate; in the example given, the data applied to both channels changes simultaneously, therefore it is immaterial which bit (the first or the second) is applied to any given channel, as long as the receiver is built so that the in-phase and quadrature channels can be recognized.

The two binary PSK modulated signals are combined; the combined signal can assume any of the four possible phases (+45°, +135°, -135°, and -45°).

It is important to note the way in which the QPSK signal is generated: two carrier signals having a 90° phase difference were separately modulated at half the data rate, and then combined. Therefore, the total spectrum width is only half the width of the spectrum pertaining to a binary PSK modulated signal carrying the same total data rate. This, of course, is the result predicted by equation (6-10), because the symbol rate of a QPSK system is only half the bit rate, but in the special case – QPSK modulation – discussed above, it is of great practical interest to interpret the QPSK signal as being the sum of two carrier signals at phase quadrature.

We note in passing that each output symbol is determined by a pair of input data bits; this pair is called a di-bit.

Another QPSK modulator is shown in Fig. 6.10; it is quite similar to that shown in Fig. 6.9, except for the fact that its output signal can assume the phases 0, +90°, 180°, and -90°. To see the practical difference between the two modulators, let us examine their output waveforms (Fig. 6.11).

The waveforms show different voltage "jumps" at the switching points, with the greater possible jump (full peak-to-peak voltage change – see di-bit VII) associated with the modulator shown in Fig. 6.10. Thus, the modulator of Fig. 6.9 is to be preferred.

190
FIGURE 6.11: Waveforms of QPSK modulators

An example of the principles involved in designing an eight-phase modulator is shown in Fig. 6.12.

FIGURE 6.12: Example of eight phase PSK modulator

The modulator consists of two QPSK modulators and one BPSK modulator, whose outputs are combined to form the eight-phase output. Vectors indicating signal phases at each sub-modulator output are also shown. Other implementations of the eight-phase modulator exist, some of them having a smaller component count.

The output symbols of the eight-phase modulator are determined by groups of three bits; these groups are called tri-bits. There are eight possible tri-bits, each represented by a specific carrier phase.

6.4.2 POWER AMPLIFIER FOR PSK TRANSMITTERS

The power amplifier operates with a constant amplitude drive, and must put out a constant-amplitude signal, therefore it need not be a linear amplifier. High efficiency Class C amplifiers are usually utilized for the final stage of a PSK transmitter.
6.4.3 TRANSMITTER OUTPUT FILTER
The transmitter output filter has two functions:

1. Reduction of transmitter harmonics,
2. Reduction of side-lobe levels.

The first function can be easily understood, since the output signal of a Class C amplifier is very rich in harmonics. This, however, is not the main problem posed to the filter designer.

The main function of the output filter is to attenuate the undesired modulation side-lobes which accompany the main lobe of a PSK signal (see Fig. 6.4, 6.5). Since the frequency bands allocated to digital transmission must be used efficiently, because of the large number of users, signals must be packed as closely as it technically feasible. The usual requirement is to allow transmitters and receivers, located at the same physical site, e.g. with antennas installed on a common antenna tower, to operate on adjacent RF channels. Although physical separation and antenna directivity contribute to a certain attenuation there remains a large power difference between the local transmitter (unfiltered) sidelobes, falling within the neighbouring receiver bandwidth, and the level of the desired signal, arriving from the distant transmitter. The main function of the transmitter's output filter is to reduce the sidelobes to a sufficiently low level that will not hamper reception of a distant transmitter signal.

In practice, transmitters used for wideband digital data transmission, i.e. for digital telephone carriers, digital television, etc. operate at fixed frequencies, allocated to them during the early stages of system design. Therefore, the output filter can usually be accurately tuned to the required frequency and can be certified by measurement, to provide the intended side-lobe attenuation.

In such cases, the designer is faced with the problem of determining the maximum amount of filtration that will not interfere with proper system operation, that is, will not degrade the radiated signal significantly.

As it is well known, filters introduce intersymbol interference (ISI), which may cause errors at the receiver even if the noise is very small.

The amount of ISI generally increases when filter bandwidth is reduced; however, for a given bandwidth, the ISI depends on the specific filter type – Butterworth, Chebyshev, etc. The sharper the cut-off characteristic, the more ISI will be introduced by the filter. The main design problem is then to select the filter providing the required attenuation slope without causing excessive ISI; an additional design constraint is to have low losses in the filters, since filter losses reduce the radiated power. The optimum filter design varies according to the intended equipment application, its operating frequency range, cost limits, etc. However, there are a few common features, such as minimum filter bandwidth; the minimum acceptable bandwidth for the output filter of a PSK transmitter is generally taken as twice the symbol rate. This allows the main lobe (see Fig. 6.4) to pass with little attenuation, therefore the amount of ISI generated by the filter (assuming Butterworth of even Chebyshev characteristics) will be acceptable. The filter is also required to have relatively steep slopes, thus a number of filter sections is generally used. To minimize losses, high-quality 'stable' low-loss components must be used in the construction of this filter.
6.5 OFFSET QPSK

QPSK is often selected when PSK transmission is desired, because it is relatively simple to implement, yet has very good error performance. However, the sidelobes in the spectrum of a QPSK signal are very strong, and as explained above, this necessitates the use of very sophisticated filters.

In order to reduce the complexity of the output filter, it would be very desirable to have a spectrum whose slope (the rate of decrease in the peaks of the individual side lobes) is steep. Ordinary PSK has a slope proportional to $1/f^2$; this slope, which is the worst of all modulation classes, is caused by the large phase jumps in the modulated waveform (see Fig. 6.12), for example, in ordinary QPSK, the maximum phase jump is $180^\circ$. If the phase jumps could be reduced, the spectrum slope would be steeper. However, reducing the phase jumps must not affect the error rate performance of this modulation class.

To see how the phase jumps can be reduced, we refer back to the QPSK modulator shown in Fig. 6.9. At the modulator’s input, a shift register is used to hold the data bits for two clock periods. After two clock periods, the shift register is loaded with two other data bits; both channels receive the new data simultaneously, so that the maximum possible phase change is $180^\circ$.

A simple method can be used to reduce the phase jump to $90^\circ$: instead of applying new data simultaneously on both I- and Q- channels, feed data alternately to each channel. The resulting timing diagram is shown in Fig. 6.13.

![Timing diagram for ordinary and offset QPSK](image)

To generate the required signal, the input data stream is demultiplexed, bit by bit, into two data streams. Each bit is latched for two input bit periods, so as to maintain unchanged the modulator input data. The block diagram of the resulting modulator is shown in Fig. 6.14.

It is important to note that the receiver need only be able to discriminate between the in-phase and quadrature channels; once the channels are recognized, the original data stream is reconstructed by multiplexing the outputs of the two channels in the correct order, that is, by simply locating the I-channel output bit before the corresponding Q-channel bit.

The spectrum of O-QPSK has a faster decrease in slope, without paying a penalty with regard to error rate; on the contrary when practical systems are compared, the O-QPSK system is better.
6.6 BINARY PSK RECEIVERS

The function of the receiver is to extract the information contained in the received signal and reconstruct, as far as possible, the original data sequence. The usual measure of performance is the error rate vs signal to noise ratio. Two types of receivers are used – coherent detection receivers and differential (incoherent) detection receivers. Both receiver types have the same sub-systems, up to the point where a reference signal must be supplied to the detector:

1. In a coherent detection receiver, the carrier signal is recovered by phase lock loops and applied to the detector.

2. In incoherent detection receivers, the signal received during the previous symbol interval is applied to the detector as a reference for evaluating the following symbol.

In general, coherent receivers have better performance; however, there are special cases where the average phase of the received signal phase varies because of irregularities in the propagation of radio signals; in such cases differential detection could perform just as well, or even better, than coherent detection.

The general block diagram of a PSK detection receiver is shown in Fig. 6.15. This block diagram is common to both types of receivers described above.

The RF signal received at the antenna is passed through an RF bandpass filter to a low-noise amplifier. The RF filter attenuates out-of-band signals generated by transmitters operating at other frequencies; its bandwidth is selected subject to the same trade-offs prevailing in the selection of the transmitter output filter (para 6.4.3. above). The filtered signal is amplified by a low-noise amplifier, to compensate for losses in the RF filter and in the mixer. The amplified signal is translated to a convenient intermediary frequency, generally 35, 70, or 140 MHz, at which the received signal is amplified and filtered. The bandwidth and characteristic of the IF filter is critical to good receiver performance, and its selection is analyzed in para. 6.7 below.

The filtered IF signal is limited, so as to remove undesirable amplitude variations. The limited signal is then applied, in parallel, to the carrier recovery circuits and to the phase detector.
In receivers utilizing coherent detection, the carrier is recovered by a phase-lock loop (PLL); in differential detection, a delay line, having a delay equal to the duration of one symbol, is used to provide the reference for the detector.

The phase detector output is filtered by a post-detection low-pass filter, whose characteristics are selected so as to eliminate as much noise as possible without impairing the regeneration of the detected signal.

The filtered signal, generally NRZ coded, is applied to the data regenerator. The regenerator and its auxiliary clock recovery circuit are similar to the circuits described in Chapter 2.

The regenerated signal is passed to a differential decoder, which cancels the differential encoding performed before modulation. Differential encoding is used even when coherent detection is used, for reasons explained in detail in para. 6.8.

We proceed now to analyze the two principal sub-systems of the receiver: the IF filter and clock recovery circuit.

### 6.7 SELECTION OF IF FILTER

The selection of an IF filter is influenced by basic requirements:

1. Elimination of noise accompanying desired signal,
2. Low distortion of signal modulation received.

Elimination of noise requires a filter with a narrow bandwidth. This contradicts the second requirement, because a filter which is too narrow will introduce high intersymbol interference (ISI).

The second requirement also involves the selection of a filter with linear phase shift (constant group delay) within its passband.
To obtain the best compromise between the two contradicting requirements set forth, the filter must have the steepest possible slopes, but must not distort the signal because of excessively non-linear phase shift.

The optimum filter characteristic, fulfilling these requirements, is the Gaussian filter; Gaussian filters or variants thereof are used in almost every digital radio equipment.

Once the filter characteristic has been determined, its bandwidth must be optimized, so as to obtain the "best performance". The problem is that optimization at a certain error rate (or signal-to-noise ratio) does not ensure the best performance at another error rate. The reason is that the optimization process balances errors due to noise (a narrower bandwidth reduces noise, and less noise means less errors) and errors due to intersymbol interference (a narrower bandwidth means increased intersymbol interference, and more errors due to this source). When balance is reached at very low signal-to-noise ratios, where most errors are due to noise, the intersymbol interference can be very large, and the error rate will not improve significantly when the receiver signal level is greater (better signal-to-noise ratio, so that errors due to ISI become dominant). On the other side, optimization at a high signal-to-noise ratio permits a large bandwidth, but more noise is accepted, and performance with low received signal levels will suffer.

Although the optimum depends on the assumed operational conditions, it does not change sharply. A typical optimization curve is shown in Fig. 6.16.

![Optimization curve for differential binary PSK filter](image)

From Fig. 6.16, it can be seen that the optimum noise bandwidth of a Gaussian filter is approximately equal to the bit rate.

The influence of IF filter bandwidth on the overall error rate is shown in Fig. 6.17. It is interesting to note that it is possible to achieve performance better than ideal coherent detection (Fig. 6.17a)
when the bandwidth is about twice the bit rate; this happens because ideal coherent detection assumes infinite bandwidth before the ideal phase detector, and therefore it accepts more noise.

![Graphs showing error rate for binary PSK as a function of S/N ratio and relative receiver filter bandwidth.](image)

**FIGURE 6.17:** Error rate for binary PSK as a function of relative receiver filter bandwidth.

When Fig. 6.16 and 6.17b are compared, it seems that they contradict each other; for an error rate of $10^{-6}$, the S/N ratio for the optimum bandwidth is seen (from Fig. 6.16) to be approximately equal to 12 dB, whereas Fig. 6.17b shows that when a bandwidth twice as large is assumed, the signal-to-noise ratio needs to be only 10 dB for the same error rate. This apparent contradiction can be resolved when the actual RF signal required to obtain the stated performance is calculated: Since noise power is directly proportional to IF filter bandwidth, the absolute noise accepted in a filter having twice the optimum bandwidth is 3 dB higher; therefore, the actual received signal must be $10 + 3 = 13$ dB above the noise level in the optimum case (with the narrower IF bandwidth). This is 1 dB higher than the signal required in the optimum case (12 dB above same noise level).

The result obtained for binary PSK transmission — Gaussian filter with bandwidth equal to bit rate — can be extended to multilevel transmission; in each case, the optimum bandwidth is nearly equal to the symbol rate.

### 6.8 CARRIER RECOVERY

Carrier recovery, or in the more general case, the problem of obtaining a reference signal for the phase detector, can be solved in three ways:

1. Differential detection,
2. Recovery of the reference signal from a pilot tone transmitted together with the modulated carrier (transmitted reference method),
3. Recovery of carrier from the modulated carrier (self-synchronized receiver).
In the following paragraphs, we will present typical implementation and the characteristics of each method.

### 6.8.1 Reference Generation for Differential Detection

The reference is generated by taking a sample of the IF signal and delaying it by one symbol interval (for binary PSK — one bit interval). The delay must be accurately set so as to be an integer multiple of the IF signal period. The resulting detector implementation is shown in Fig. 6.18.

![Differential detector implementation](image)

The delay line is usually a piece of high quality coaxial cable; the main requirements are low loss and high stability. Rigid coaxial cable with Teflon insulation is generally used; its delay constant is approximately 20 cm/nsec. The small delay makes this type of delay device impractical when long delays (low bit rates) are required, so other types of delay lines are also used, generally consisting of discrete components.

An emerging technology, surface acoustic wave (SAW) filters, is also capable of providing high quality, long delay elements; SAW devices may be practical down to transmission rates of several megabits/second.

### 6.8.2 Transmitted Reference Systems

The transmitted reference system utilizes a pilot tone, which is transmitted along with the information carrying signal. Generally, the pilot tone is generated by allowing "leakage" of the unmodulated carrier signal through the phase modulator, but other convenient frequencies within the intended IF reception bandwidth could be used as well.

At the receiver, a sample of the IF signal is taken to a PLL tuned to the pilot frequency. The PLL serves as a narrow-band tuned filter, so that the output signal will be relatively clean and free of noise even when the received signal is near reception threshold. The filtered pilot signal is applied to the reference input of a frequency synthesizer which regenerates the carrier signal to be applied to the phase detector. The block diagram of this carrier recovery system is shown in Fig. 6.19.

![Generation of carrier reference by transmitted reference system](image)

The transmitted reference system does allow for simple implementation, but it has one important disadvantage; since the total transmitter power available is constant, this power must be divided between the information-carrying signal and the pilot tone. There is no single optimum division, because the characteristics of the pilot tone recovery PLL depend on the intended system application and operating conditions.
For example, when fast synchronization is desired, e.g. for mobile radio equipment operating on a push-to-talk basis, the PLL must have a relatively wide bandwidth to lock on rapidly. To minimize noise on the recovered signal a certain pilot signal/noise ratio must be provided; therefore, when the PLL bandwidth is increased, a larger pilot signal must be transmitted to obtain the required signal/noise. Therefore, less power remains for transmission of information. Since errors are caused by both noise and phase jitter on the reference signal, an optimum partition of available transmitter power must be sought, which will minimize the total error rate. The optimum partition reached for the application given above will differ from the partition that will be found optimal for a system which transmits continuously, so that lock-on time is not important.

Experience shows that in most cases, the performance of a transmitted reference system is inferior to that of a self-synchronized system, optimized under similar conditions.

6.8.3 RECOVERY OF CARRIER FROM THE MODULATED SIGNAL

6.8.3.1 Squaring Loop Carrier Recovery System
In order to recover the carrier signal, it is generally necessary to remove the modulation. For the binary PSK case, this can be easily achieved by doubling the frequency of the input signal. Since the phase changes in the PSK signal are either 0, or ± 180°, doubling results in 0° or ± 360° changes, that is, phase modulation is eliminated. The effect of doubling is shown graphically in Fig. 6.20.

![Figure 6.20: Operation of squarer](image)

Although ideally a squarer should be used, a simple full-wave rectifier can also fulfill the same function; this is shown in Fig. 6.20. The only difference is that the rectifier output has a large harmonic content, but since a narrow bandwidth PLL is anyhow connected to its output, this has little effect on the output signal. The typical block diagram of the squaring carrier recovery system is shown in Fig. 6.21.

![Figure 6.21: Squaring loop carrier recovery system](image)
The system consists of a rectifier, or squarer, followed by a PLL which generates a clean, stable signal at twice the input frequency. The PLL output signal is divided by 2, so as to recover the original frequency, and the divided signal is applied, as a reference, to the phase detector.

Because of the phase doubling, it is impossible to establish an unambiguous phase relationship between the input and output signal; that is, the output signal can lock, with equal probability, at 0 or 180° with reference to the input signal. A phase shift of 180° means that the detector will put out a "0" when a "1" has been transmitted, and vice-versa.

The only way to remove the ambiguity is to transmit a known training sequence after each re-locking, which will enable selection of the proper reference output phase. In most cases, this is not possible. Instead, differential encoding is applied to the transmitted data, so that information is carried by transitions, no matter in which direction. This obviates the need to establish the correct phase relationship between the recovered carrier and the IF signal.

The main design parameters of the carrier recovery system are its acquisition time and the jitter as a function of input signal/noise ratio.

Jitter is reduced by having a loop low pass filter with a low cut-off frequency; this results in long acquisition times. To shorten the acquisition time, two filter cut-off frequencies are sometimes used, one during normal operation, the other, higher than the normal cut-off, during the acquisition process.

The influence of jitter on signal detection is similar to the influence of jitter on the detection of base band signals (see Chapter 2). Jitter in excess of a few degrees (peak-to-peak) reduces the effective signal power; typically, 1 dB is lost in binary PSK systems when the jitter reaches 26° peak-to-peak.

However, not only jitter but also a static phase error causes degradation; to reduce the static error, a high loop gain is required, together with an adjustment for the PLL output phase that will permit cancellation of the static error.

6.8.3.2 Costas Loop for Carrier Recovery
The PSK signal is a suppressed-carrier signal, therefore it is not possible to lock directly to the carrier. Non-linear processing is required in order to obtain information sufficient for regenerating the carrier from the modulated signal. This has been already shown for the squaring loop.

However, there exists a special type of phase lock loop, intended for tracking suppressed carrier signals; this is the Costas loop (Fig. 6.22).

Unlike an ordinary PLL, the Costas loop utilizes the signal sidebands, instead of the carrier, for extracting the tracking signal; in fact, for a given total signal power, the Costas loop operates best when all power is allocated to the sidebands, yet it can accurately regenerate the missing carrier.

The Costas loop has two phase detectors (mixers) one fed with a signal in-phase with the VCO signal, another with a signal in quadrature with the VCO signal: the outputs of the two phase detectors are filtered by two identical low pass filters, termed the arm filters. The bandwidth of these filters is made sufficiently wide to allow the modulation signal to pass without attenuation, but the filters remove the harmonics of the input signal. The filtered arm signals are applied to a multiplier, which generates the error voltage for the control of the VCO. This signal is filtered by another low-pass filter, the loop filter, whose transfer function determines the dynamic characteristics of the Costas PLL.
To obtain a DC control voltage for the VCO, in-phase signal components must exist at the outputs of both arm filters, these components stem from modulation sidebands, not from the carrier signal. Therefore, the Costas loop can track suppressed carrier signals, but will lose lock if modulation stops.

The overall performance of the Costas loop is similar to that of a squarer followed by a PLL, except that it is capable of better tracking.

6.9 PHASE DETECTOR AND POST DETECTION FILTER

The implementation of the phase detector used in a PSK receiver depends on the IF frequency and transmission rate. For frequencies up to several tens of megacycles, active multiplier circuits are most effective, whereas at higher frequencies, double balanced diode mixers are normally used.

The output of the phase detector is filtered by a post-detection low-pass filter. This filter performs two functions:

1. Removes harmonics of the IF signal. This usually does not pose any problem, since the IF frequency is orders of magnitude higher than the highest modulation frequency.
2. Filters the demodulated signal, to remove noise outside the range of useful modulation, to improve the signal-to-noise ratio seen by the data regenerator.

From a theoretical point-of-view, it is very difficult to include the effects of the post detection filter in the error rate calculation. The filter characteristics are therefore selected experimentally, following a trade-off similar to that made in the selection of the IF filter bandwidth (para 6.7 above).

6.10 CALCULATION OF PSK RECEIVER PERFORMANCE

The principles of performance calculation were explained in Chapter 5, paragraph 5.8. The same procedure is also applicable to PSK receivers, namely:

\[ \text{Figure 6.22: Costas PLL block diagram} \]
1. Calculation of signal-to-noise ratio seen by the detector,
2. Calculation of detector's error rate at the signal-to-noise ratio found in step 1.

Additional parameters of interest may include carrier recovery system lock time and output filter versus signal level. These parameters are obtained during the design of the carrier recovery system as a function of the signal-to-noise ratio at the input of this system; using the results of step 1, the parameters can be referenced to the actual received signal level.

6.11 QUADRATURE PSK RECEIVER

The design and implementation of a QPSK receiver is very similar to that of a binary PSK system, except for the IF filter bandwidth and the details of carrier recovery and data demodulation.

The optimum value of the IF filter for a QPSK receiver is slightly over half the bit rate (or slightly over the symbol rate); the optimization curve for a Gaussian IF filter is shown in Fig. 6.23.

![Graph showing S/N vs. Normalized Bandwidth for QPSK receiver]

**FIGURE 6.23:**
Optimization curve for the IF bandwidth of a QPSK receiver employing differential detection

The corresponding error rate curves are shown in Fig. 6.24.

It is interesting that, theoretically, QPSK has performance identical to binary PSK but permits reduction of the transmission bandwidth to half that required by binary PSK.

This conclusion is based on the implementation of QPSK as two independent binary PSK channels, modulated on carriers of the same frequency but in quadrature. This implementation has been
described in paragraph 6.5. Since the bit rate of each QPSK channel is half that of the binary PSK system, the required receiver bandwidth is also halved. Therefore, for a given received signal strength, the signal/noise ratio seen by the detector of each QPSK channel would be 3 dB higher than that seen by the detector of the binary system.

**FIGURE 6.24:**
*Error rate for differentially detected QPSK as a function of the relative receiver filter bandwidth*

To make a fair comparison, we must assume equal transmitter power for both QPSK and BPSK systems. Considering the QPSK transmitter, we must divide its output power equally between the two channels, thus each channel receives 3 dB less power than the BPSK channels. This means that the signal-to-noise ratio seen by the detector of each QPSK channel equals that seen by the detector of the BPSK channel, and they will therefore have identical error rates.

The equivalence of BPSK and QPSK explains the wide use of QPSK for data transmission. However, in practice, degradations caused by the carrier recovery system and imperfections in the IF filter characteristics are slightly larger for the QPSK system, but the difference is not critical.

In order to use the advantages of QPSK, the detection system should use coherent detection and be composed of two binary channels, as shown in Fig. 6.25.

**FIGURE 6.25:** *QPSK detector and data regenerator*
The detector shown in Fig. 6.25 has two channels, each operating in the binary mode. Each channel consists of a phase detector followed by a low-pass filter, the filter is optimized for transmission at half the total data rate. The filtered output signal of each detector is applied to a NRZ data regenerator and also to the clock recovery circuits. The regenerated data stream passes through a differential decoder, which removes the effects of the differential encoder used at the transmitter. The decoded data streams of the two channels are multiplexed bit-by-bit, to reverse the effect of the demultiplexer used on the modulator (Fig. 6.9, 6.10). The details of the multiplexing operation depend on the modulation type; for ordinary QPSK, the two bits appear simultaneously at the differential decoder outputs, whereas, for offset QPSK, the output bit from the quadrature channel appears after the in-phase channel bit (see Fig. 6.13).

The differential decoder is shown in Fig. 6.26.

![Differential decoder](image)

**FIGURE 6.26:**
Differential decoder

(b) Typical waveforms

The decoding rule is to insert "1"s into bit slots which contain transitions; in the remaining slots, "0"s are inserted. The typical waveforms shown in Fig. 6.26b demonstrate the reversal of the encoding operation shown in Fig. 6.7b.

Differential decoding has a serious flaw; if an error occurs, it will also cause an error in the decoding of the following bit (unless this bit is also erroneous) and therefore error doubling occurs. This introduces an inevitable degradation in performance. However, despite the fact that the error rate is doubled, the equivalent degradation is a few tenths of a dB, i.e. it can be offset by increasing the received signal level by several tenths of a dB.

### 6.12 CARRIER RECOVERY FOR QPSK RECEIVERS

The three alternatives for the supply of reference signals to the detector, described in para 6.8 with respect to the binary PSK receiver, also apply with slight modifications to QPSK receivers. The following paragraphs deal with the modifications required in order to permit recovery of the reference in QPSK receivers.
6.12.1 Fourth-Power Loop

Since the QPSK signal can assume four different phases which are multiples of 90°, squaring (doubling of the input frequency) is not sufficient to remove the modulation. To achieve this, the frequency of the input signal must be quadrupled, that is, raised to the fourth power. The fourth harmonic of the quadrupler is then filtered by a narrow bandwidth PLL and divided by four, to re-establish the original frequency. The resulting block diagram is shown in Fig. 6.27.

![Diagram of Fourth-Power QPSK Carrier Recovery System](image)

**Figure 6.27:** Fourth power QPSK carrier recovery system

The divider is designed so as to provide four outputs, having a 90° phase difference between them. Two of these outputs are taken to the phase detectors (Fig. 6.25), however, it is not possible to predetermine the correct outputs, because of the fourfold phase ambiguity introduced by the frequency quadrupling.

The selection can be made by transmitting a known training sequence following each re-locking of the carrier recovery system, so that the proper polarity is established. However, when differential coding is used, the problem is somewhat reduced, because the selection of a wrong pair of outputs will only lead to inversion between the in-phase and quadrature channels, without affecting the data transmitted on each channel.

![Diagram of Four-Phase Costas PLL](image)

**Figure 6.28:** Four-phase Costas PLL
6.12.2 FOUR-PHASE COSTAS LOOP

The Costas PLL can be modified to allow operation with four-phase signals (or even more). This requires two groups of mixers (two mixers in each group), one for each of the two quadrature signals composing the QPSK signals.

The signals of each pair of channels are multiplied together, until the control signal for the voltage-controlled oscillator is obtained. The control signal is filtered by the loop filter, whose characteristics determine the stability and bandwidth of the PLL. In addition to this filter, four low-pass filters are connected at the outputs of the mixers; the bandwidth of these filters is wide enough to pass the modulation signals, but sufficiently low in relation to the harmonics of the IF signal so that these are strongly attenuated.

6.13 EIGHT AND SIXTEEN-PHASE PSK SYSTEMS

To obtain the high bandwidth efficiency required by wideband commercial microwave digital radio, it becomes necessary to use 8- and even 16- phase PSK. The implementation of these systems is quite similar to the QPSK system. In these systems, a significant range reduction is incurred because of the high-order modulation and because of the increased influence of imperfections and tolerances in the radio equipment circuits. However, this is usually not critical, because commercial links can be engineered to tolerate degradation and poor propagation conditions.