6

CARRIER AND SYMBOL SYNCHRONIZATION

We have observed that in a digital communication system, the output of the demodulator must be sampled periodically, once per symbol interval, in order to recover the transmitted information. Since the propagation delay from the transmitter to the receiver is generally unknown at the receiver, symbol timing must be derived from the received signal in order to synchronously sample the output of the demodulator.

The propagation delay in the transmitted signal also results in a carrier offset, which must be estimated at the receiver if the detector is phase-coherent. In this chapter, we consider methods for deriving carrier and symbol synchronization at the receiver.

6-1 SIGNAL PARAMETER ESTIMATION

Let us begin by developing a mathematical model for the signal at the input to the receiver. We assume that the channel delays the signals transmitted through it and corrupts them by the addition of gaussian noise. Hence, the received signal may be expressed as

$$r(t) = s(t - \tau) + n(t)$$

where

$$s(t) = \text{Re}\,[s_l(t)e^{j2\pi f_c t}] \tag{6-1-1}$$

and where τ is the propagation delay and $s_l(t)$ is the equivalent lowpass signal.

The received signal may be expressed as

$$r(t) = \text{Re}\,\{[s_l(t - \tau)e^{j\phi} + z(t)]e^{j2\pi f_c t}\} \tag{6-1-2}$$

where the carrier phase ϕ, due to the propagation delay τ, is $\phi = -2\pi f_c \tau$.

Now, from this formulation, it may appear that there is only one signal parameter to be estimated, namely, the propagation delay, since one can determine ϕ from knowledge of f_c and τ. However, this is not the case. First of all, the oscillator that generates the carrier signal for demodulation at the receiver is generally not synchronous in phase with that at the transmitter. Furthermore, the two oscillators may be drifting slowly with time, perhaps in different directions. Consequently, the received carrier phase is not only dependent on the time delay τ. Furthermore, the precision to which one must synchronize in time for purpose of demodulating the received signal depends on the symbol interval T. Usually, the estimation error in estimating τ must be a relatively small fraction of T. For example, $\pm 1\%$ of T is adequate for practical applications. However, this level of precision is generally inadequate for estimating the carrier phase, even if ϕ depends only on τ. This is due to the fact that f_c is generally large, and, hence, a small estimation error in τ causes a large phase error.

In effect, we must estimate both parameters τ and ϕ in order to demodulate and coherently detect the received signal. Hence, we may express the received signal as

$$r(t) = s(t; \phi, \tau) + n(t) \tag{6-1-3}$$

where ϕ and τ represent the signal parameters to be estimated. To simplify the notation, we let ψ denote the parameter vector $\{\phi, \tau\}$, so that $s(t; \phi, \tau)$ is simply denoted by $s(t; \psi)$.

There are basically two criteria that are widely applied to signal parameter estimation: the *maximum-likelihood* (ML) criterion and the *maximum* a posteriori *probability* (MAP) criterion. In the MAP criterion, the signal parameter vector ψ is modeled as random, and characterized by an a priori probability density function $p(\psi)$. In the maximum-likelihood criterion, the signal parameter vector ψ is treated as deterministic but unknown.

By performing an orthonormal expansion of $r(t)$ using N orthonormal functions $\{f_n(t)\}$, we may represent $r(t)$ by the vector of coefficients $[r_1\ r_2\ \ldots\ r_N] \equiv \mathbf{r}$. The joint pdf of the random variables $[r_1\ r_2\ \ldots\ r_N]$ in the expansion can be expressed as $p(\mathbf{r} \mid \psi)$. Then, the ML estimate of ψ is the value that maximizes $p(\mathbf{r} \mid \psi)$. On the other hand, the MAP estimate is the value of ψ that maximizes the a posteriori probability density function

$$p(\psi \mid \mathbf{r}) = \frac{p(\mathbf{r} \mid \psi)p(\psi)}{p(\mathbf{r})} \tag{6-1-4}$$

We note that if there is no prior knowledge of the parameter vector ψ, we may assume that $p(\psi)$ is uniform (constant) over the range of values of the parameters. In such a case, the value of ψ that maximizes $p(\mathbf{r} \mid \psi)$ also maximizes $p(\psi \mid \mathbf{r})$. Therefore, the MAP and ML estimates are identical.

In our treatment of parameter estimation given below, we view the parameters ϕ and τ as unknown, but deterministic. Hence, we adopt the ML criterion for estimating them.

In the ML estimation of signal parameters, we require that the receiver extract the estimate by observing the received signal over a time interval $T_0 \geq T$, which is called the observation interval. Estimates obtained from a single observation interval are sometimes called one-shot estimates. In practice, however, the estimation is performed on a continuous basis by using tracking loops (either analog or digital) that continuously update the estimates. Nevertheless, one-shot estimates yield insight for tracking loop implementation. In addition, they prove useful in the analysis of the performance of ML estimation, and their performance can be related to that obtained with a tracking loop.

6-1-1 The Likelihood Function

Although it is possible to derive the parameter estimates based on the joint pdf of the random variables $[r_1 \ r_2 \ \ldots \ r_N]$ obtained from the expansion of $r(t)$, it is convenient to deal directly with the signal waveforms when estimating their parameters. Hence, we shall develop a continuous-time equivalent of the maximization of $p(\mathbf{r} \mid \boldsymbol{\psi})$.

Since the additive noise $n(t)$ is white and zero-mean gaussian, the joint pdf $p(\mathbf{r} \mid \boldsymbol{\psi})$ may be expressed as

$$p(\mathbf{r} \mid \boldsymbol{\psi}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left\{-\sum_{n=1}^{N} \frac{[r_n - s_n(\boldsymbol{\psi})]^2}{2\sigma^2}\right\} \tag{6-1-5}$$

where

$$\begin{aligned} r_n &= \int_{T_0} r(t) f_n(t)\, dt \\ s_n(\boldsymbol{\psi}) &= \int_{T_0} s(t; \boldsymbol{\psi}) f_n(t)\, dt \end{aligned} \tag{6-1-6}$$

where T_0 represents the integration interval in the expansion of $r(t)$ and $s(t; \boldsymbol{\psi})$.

We note that the argument in the exponent may be expressed in terms of the signal waveforms $r(t)$ and $s(t; \boldsymbol{\psi})$, by substituting from (6-1-6) into (6-1-5). That is,

$$\frac{1}{2\sigma^2} \sum_{n=1}^{N} [r_n - s_n(t; \boldsymbol{\psi})]^2 = \frac{1}{N_0} \int_{T_0} [r(t) - s(t; \boldsymbol{\psi})]^2\, dt \tag{6-1-7}$$

where the proof is left as an exercise for the reader (see Problem 6-1). Now, the maximization of $p(\mathbf{r} \mid \boldsymbol{\psi})$ with respect to the signal parameters $\boldsymbol{\psi}$ is equivalent to the maximization of the *likelihood function*.

$$\Lambda(\boldsymbol{\psi}) = \exp\left\{-\frac{1}{N_0} \int_{T_0} [r(t) - s(t; \boldsymbol{\psi})]^2\, dt\right\} \tag{6-1-8}$$

Below, we shall consider signal parameter estimation from the viewpoint of maximizing $\Lambda(\boldsymbol{\psi})$.

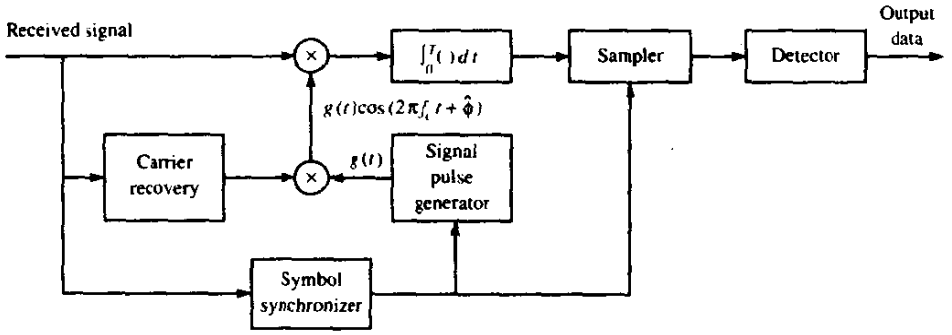


FIGURE 6-1-1 Block diagram of binary PSK receiver.

6-1-2 Carrier Recovery and Symbol Synchronization in Signal Demodulation

Symbol synchronization is required in every digital communication system which transmits information synchronously. Carrier recovery is required if the signal is detected coherently.

Figure 6-1-1 illustrates the block diagram of a binary PSK (or binary PAM) signal demodulator and detector. As shown, the carrier phase estimate $\hat{\phi}$ is used in generating the reference signal $g(t)\cos(2\pi f_c t + \hat{\phi})$ for the correlator. The symbol synchronizer controls the sampler and the output of the signal pulse generator. If the signal pulse is rectangular then the signal generator can be eliminated.

The block diagram of an M-ary PSK demodulator is shown in Fig. 6-1-2. In this case, two correlators (or matched filters) are required to correlate the received signal with the two quadrature carrier signals $g(t)\cos(2\pi f_c t + \hat{\phi})$ and $g(t)\sin(2\pi f_c t + \hat{\phi})$, where $\hat{\phi}$ is the carrier phase estimate. The detector is now a phase detector, which compares the received signal phases with the possible transmitted signal phases.

The block diagram of a PAM signal demodulator is shown in Fig. 6-1-3. In this case, a single correlator is required, and the detector is an amplitude detector, which compares the received signal amplitude with the possible transmitted signal amplitudes. Note that we have included an automatic gain control (AGC) at the front-end of the demodulator to eliminate channel gain variations, which would affect the amplitude detector. The AGC has a relatively long time constant, so that it does not respond to the signal amplitude variations that occur on a symbol-by-symbol basis. Instead, the AGC maintains a fixed average (signal plus noise) power at its output.

Finally, we illustrate the block diagram of a QAM demodulator in Fig. 6-1-4. As in the case of PAM, an AGC is required to maintain a constant average power signal at the input to the demodulator. We observe that the demodulator is similar to a PSK demodulator, in that both generate in-phase and quadrature signal samples (X, Y) for the detector. In the case of QAM,

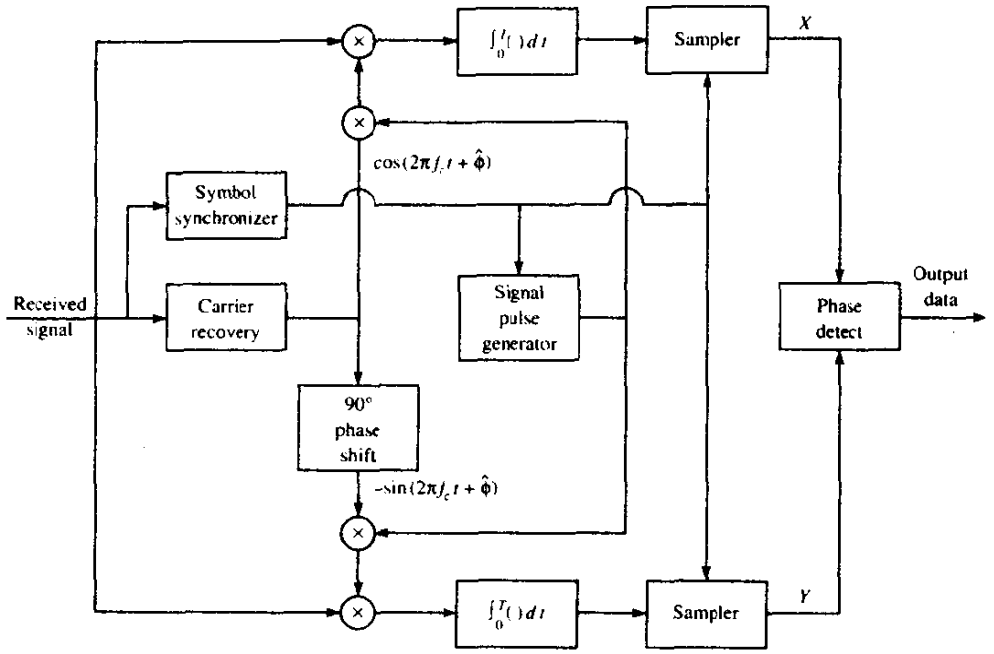


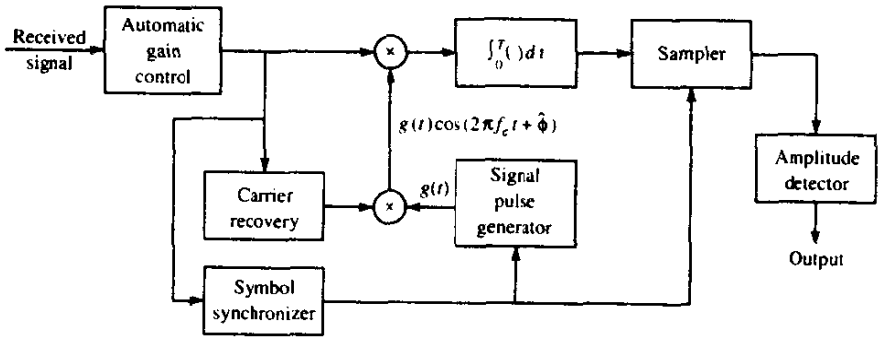
FIGURE 6-1-2 Block diagram of M-ary PSK receiver.

the detector computes the euclidean distance between the received noise-corrupted signal point and the M possible transmitted points, and selects the signal closest to the received point.

6-2 CARRIER PHASE ESTIMATION

There are two basic approaches for dealing with carrier synchronization at the receiver. One is to multiplex, usually in frequency, a special signal, called a pilot signal, that allows the receiver to extract and, thus, to synchronize its local oscillator to the carrier frequency and phase of the received signal. When

FIGURE 6-1-3 Block diagram of M-ary PAM receiver.



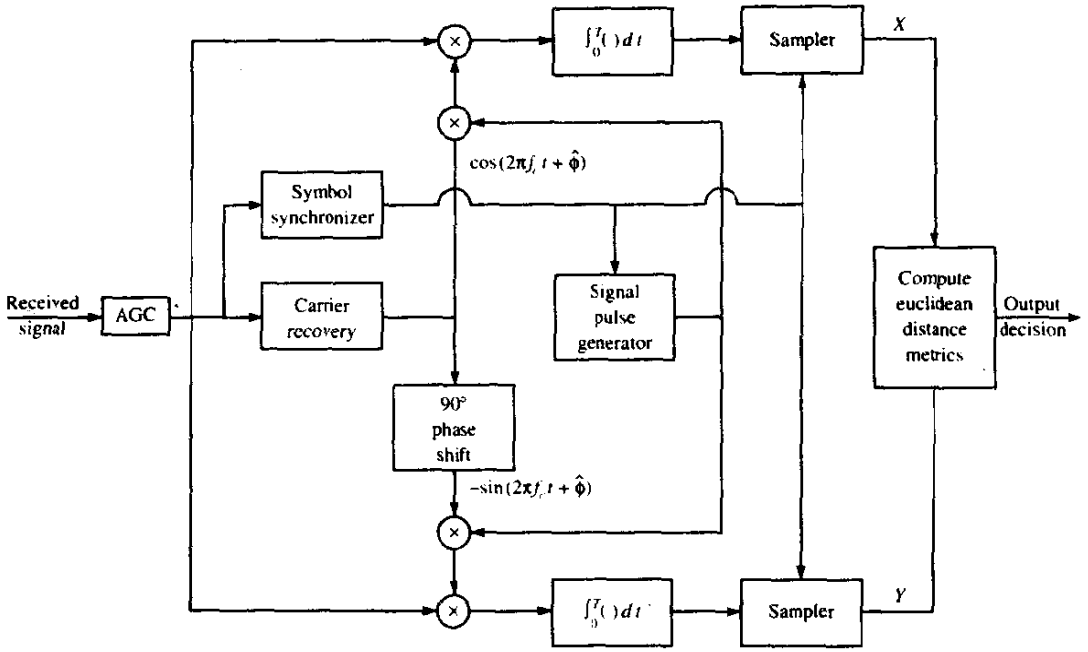


FIGURE 6-1-4 Block diagram of QAM receiver.

an unmodulated carrier component is transmitted along with the information-bearing signal, the receiver employs a phase-locked loop (PLL) to acquire and track the carrier component. The PLL is designed to have a narrow bandwidth so that it is not significantly affected by the presence of frequency components from the information-bearing signal.

The second approach, which appears to be more prevalent in practice, is to derive the carrier phase estimate directly from the modulated signal. This approach has the distinct advantage that the total transmitter power is allocated to the transmission of the information-bearing signal. In our treatment of carrier recovery, we confine our attention to the second approach; hence, we assume that the signal is transmitted via suppressed carrier.

In order to emphasize the importance of extracting an accurate phase estimate, let us consider the effect of a carrier phase error on the demodulation of a double-sideband, suppressed carrier (DSB/SC) signal. To be specific, suppose we have an amplitude-modulated signal of the form

$$s(t) = A(t)\cos(2\pi f_c t + \phi) \tag{6-2-1}$$

If we demodulate the signal by multiplying $s(t)$ with the carrier reference

$$c(t) = \cos(2\pi f_c t + \hat{\phi}) \tag{6-2-2}$$

we obtain

$$c(t)s(t) = \tfrac{1}{2}A(t)\cos(\phi - \hat{\phi}) + \tfrac{1}{2}A(t)\cos(4\pi f_c t + \phi + \hat{\phi})$$

The double-frequency component may be removed by passing the product signal $c(t)s(t)$ through a lowpass filter. This filtering yields the information-bearing signal

$$y(t) = \tfrac{1}{2}A(t)\cos(\phi - \hat{\phi}) \tag{6-2-3}$$

Note that the effect of the phase error $\phi - \hat{\phi}$ is to reduce the signal level in voltage by a factor $\cos(\phi - \hat{\phi})$ and in power by a factor $\cos^2(\phi - \hat{\phi})$. Hence, a phase error of 10° results in a signal power loss of 0.13 dB, and a phase error of 30° results in a signal power loss of 1.25 dB in an amplitude-modulated signal.

The effect of carrier phase errors in QAM and multiphase PSK is much more severe. The QAM and M-PSK signals may be represented as

$$s(t) = A(t)\cos(2\pi f_c t + \phi) - B(t)\sin(2\pi f_c t + \phi) \tag{6-2-4}$$

This signal is demodulated by the two quadrature carriers

$$\begin{aligned} c_c(t) &= \cos(2\pi f_c t + \hat{\phi}) \\ c_s(t) &= -\sin(2\pi f_c t + \hat{\phi}) \end{aligned} \tag{6-2-5}$$

Multiplication of $s(t)$ with $c_c(t)$ followed by lowpass filtering yields the in-phase component

$$y_I(t) = \tfrac{1}{2}A(t)\cos(\phi - \hat{\phi}) - \tfrac{1}{2}B(t)\sin(\phi - \hat{\phi}) \tag{6-2-6}$$

Similarly, multiplication of $s(t)$ by $c_s(t)$ followed by lowpass filtering yields the quadrature component

$$y_Q(t) = \tfrac{1}{2}B(t)\cos(\phi - \hat{\phi}) + \tfrac{1}{2}A(t)\sin(\phi - \hat{\phi}) \tag{6-2-7}$$

The expressions (6-2-6) and (6-2-7) clearly indicate that the phase error in the demodulation of QAM and M-PSK signals has a much more severe effect than in the demodulation of a PAM signal. Not only is there a reduction in the power of the desired signal component by a factor $\cos^2(\phi - \hat{\phi})$, but there is also crosstalk interference from the in-phase and quadrature components. Since the average power levels of $A(t)$ and $B(t)$ are similar, a small phase error causes a large degradation in performance. Hence, the phase accuracy requirements for QAM and multiphase coherent PSK are much higher than DSB/SC PAM.

6-2-1 Maximum-Likelihood Carrier Phase Estimation

First, we derive the maximum-likelihood carrier phase estimate. For simplicity, we assume that the delay τ is known and, in particular, we set $\tau = 0$. The function to be maximized is the likelihood function given in (6-1-8). With ϕ substituted for $\boldsymbol{\psi}$, this function becomes

$$\begin{aligned} \Lambda(\phi) &= \exp\left\{-\frac{1}{N_0}\int_{T_0}[r(t) - s(t;\phi)]^2\,dt\right\} \\ &= \exp\left\{-\frac{1}{N_0}\int_{T_0} r^2(t)\,dt + \frac{2}{N_0}\int_{T_0} r(t)s(t;\phi)\,dt - \frac{1}{N_0}\int_{T_0} s^2(t;\phi)\,dt\right\}. \end{aligned} \tag{6-2-8}$$

Note that the first term of the exponential factor does not involve the signal parameter ϕ. The third term, which contains the integral of $s^2(t;\phi)$, is a constant equal to the signal energy over the observation interval T_0 for any value of ϕ. Only the second term, which involves the cross-correlation of the received signal $r(t)$ with the signal $s(t;\phi)$, depends on the choice of ϕ. Therefore, the likelihood function $\Lambda(\phi)$ may be expressed as

$$\Lambda(\phi) = C \exp\left[\frac{2}{N_0}\int_{T_0} r(t)s(t;\phi)\,dt\right] \tag{6-2-9}$$

where C is a constant independent of ϕ.

The ML estimate $\hat{\phi}_{ML}$ is the value of ϕ that maximizes $\Lambda(\phi)$ in (6-2-9). Equivalently, the value $\hat{\phi}_{ML}$ also maximizes the logarithm of $\Lambda(\phi)$, i.e., the log-likelihood function

$$\Lambda_L(\phi) = \frac{2}{N_0}\int_{T_0} r(t)s(t;\phi)\,dt \tag{6-2-10}$$

Note that in defining $\Lambda_L(\phi)$ we have ignored the constant term $\ln C$.

Example 6-2-1

As an example of the optimization to determine the carrier phase, let us consider the transmission of the unmodulated carrier $A\cos 2\pi f_c t$. The received signal is

$$r(t) = A\cos(2\pi f_c t + \phi) + n(t)$$

where ϕ is the unknown phase. We seek the value ϕ, say $\hat{\phi}_{ML}$, that maximizes

$$\Lambda_L(\phi) = \frac{2A}{N_0}\int_{T_0} r(t)\cos(2\pi f_c t + \phi)\,dt$$

A necessary condition for a maximum is that

$$\frac{d\Lambda_L(\phi)}{d\phi} = 0$$

This condition yields

$$\int_{T_0} r(t)\sin(2\pi f_c t + \hat{\phi}_{ML})\,dt = 0 \tag{6-2-11}$$

or, equivalently,

$$\hat{\phi}_{ML} = -\tan^{-1}\left[\int_{T_0} r(t)\sin 2\pi f_c t\,dt \Big/ \int_{T_0} r(t)\cos 2\pi f_c t\,dt\right] \tag{6-2-12}$$

We observe that the optimality condition given by (6-2-11) implies the use

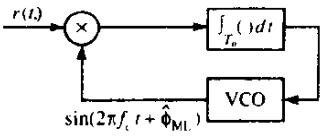


FIGURE 6-2-1 A PLL for obtaining the ML estimate of the phase of an unmodulated carrier.

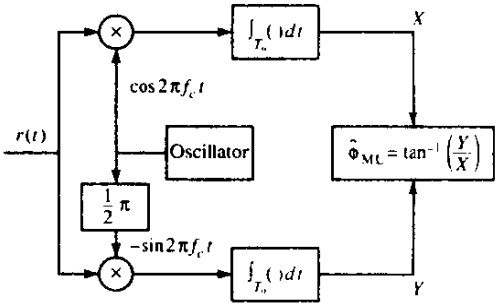


FIGURE 6-2-2 A (one-shot) ML estimate of the phase of an unmodulated carrier.

of a loop to extract the estimate as illustrated in Fig. 6-2-1. The loop filter is an integrator whose bandwidth is proportional to the reciprocal of the integration interval T_0. On the other hand, (6-2-12) implies an implementation that uses quadrature carriers to cross-correlate with $r(t)$. Then, $\hat{\phi}_{ML}$ is the inverse tangent of the ratio of these two correlator outputs, as shown in Fig. 6-2-2. Note that this estimation scheme yields $\hat{\phi}_{ML}$ explicitly.

This example clearly demonstrates that the PLL provides the ML estimate of the phase of an unmodulated carrier.

6-2-2 The Phase-Locked Loop

The PLL basically consists of a multiplier, a loop filter, and a voltage-controlled oscillator (VCO), as shown in Fig. 6-2-3. If we assume that the input to the PLL is the sinusoid $\cos(2\pi f_c t + \phi)$ and the output of the VCO is $\sin(2\pi f_c t + \hat{\phi})$, where $\hat{\phi}$ represents the estimate of ϕ, the product of these two signals is

$$\begin{aligned} e(t) &= \cos(2\pi f_c t + \phi)\sin(2\pi f_c t + \hat{\phi}) \\ &= \tfrac{1}{2}\sin(\hat{\phi} - \phi) + \tfrac{1}{2}\sin(4\pi f_c t + \phi + \hat{\phi}) \end{aligned} \tag{6-2-13}$$

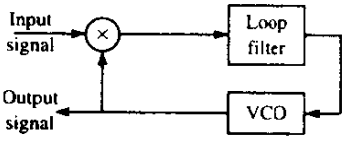


FIGURE 6-2-3 Basic elements of a phase-located loop (PLL).

The loop filter is a lowpass filter that responds only to the low-frequency component $\frac{1}{2}\sin(\hat{\phi} - \phi)$ and removes the component at $2f_c$. This filter is usually selected to have the relatively simple transfer function

$$G(s) = \frac{1 + \tau_2 s}{1 + \tau_1 s} \tag{6-2-14}$$

where τ_1 and τ_2 are design parameters ($\tau_1 \gg \tau_2$) that control the bandwidth of the loop. A higher-order filter that contains additional poles may be used if necessary to obtain a better loop response.

The output of the loop filter provides the control voltage $v(t)$ for the VCO. The VCO is basically a sinusoidal signal generator with an instantaneous phase given by

$$2\pi f_c t + \hat{\phi}(t) = 2\pi f_c t + K \int_{-\infty}^{t} v(\tau)\, d\tau \tag{6-2-15}$$

where K is a gain constant in rad/V. Hence,

$$\hat{\phi}(t) = K \int_{-\infty}^{t} v(\tau)\, d\tau \tag{6-2-16}$$

By neglecting the double-frequency term resulting from the multiplication of the input signal with the output of the VCO, we may reduce the PLL into the equivalent closed-loop system model shown in Fig. 6-2-4. The sine function of the phase difference $\phi - \hat{\phi}$ makes this system nonlinear, and, as a consequence, the analysis of its performance in the presence of noise is somewhat involved but, nevertheless, it is mathematically tractable for some simple loop filters.

In normal operation when the loop is tracking the phase of the incoming carrier, the phase error $\phi - \hat{\phi}$ is small and, hence,

$$\sin(\hat{\phi} - \phi) \approx \hat{\phi} - \phi \tag{6-2-17}$$

With this approximation, the PLL becomes linear and is characterized by the closed-loop transfer function

$$H(s) = \frac{KG(s)/s}{1 + KG(s)/s} \tag{6-2-18}$$

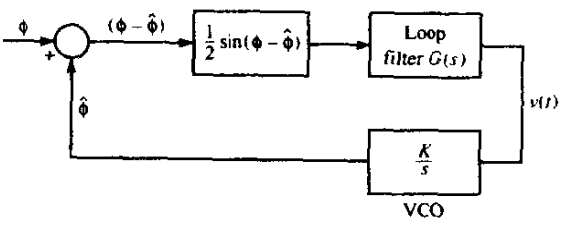


FIGURE 6-2-4 Model of phase-locked loop.

where the factor of $\frac{1}{2}$ has been absorbed into the gain parameter K. By substituting from (6-2-14) for $G(s)$ into (6-2-18), we obtain

$$H(s) = \frac{1 + \tau_2 s}{1 + (\tau_2 + 1/K)s + (\tau_1/K)s^2} \tag{6-2-19}$$

Hence, the closed-loop system for the linearized PLL is second-order when $G(s)$ is given by (6-2-14). The parameter τ_2 controls the position of the zero, while K and τ_1 are used to control the position of the closed-loop system poles. It is customary to express the denominator of $H(S)$ in the standard form

$$D(s) = s^2 + 2\zeta\omega_n s + \omega_n^2 \tag{6-2-20}$$

where ζ is called the *loop damping factor* and ω_n is the natural frequency of the loop. In terms of the loop parameters, $\omega_n = \sqrt{K/\tau_1}$, and $\zeta = (\tau_2 + 1/K)/2\omega_n$, the closed-loop transfer function becomes

$$H(s) = \frac{(2\zeta\omega_n - \omega_n^2/K)s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{6-2-21}$$

The (one-sided) noise-equivalent bandwidth (see Problem 2-24) of the loop is

$$\begin{aligned} B_{eq} &= \frac{\tau_2^2(1/\tau_2^2 + K/\tau_1)}{4(\tau_2 + 1/K)} \\ &= \frac{1 + (\tau_2\omega_n)^2}{8\zeta\omega_n} \end{aligned} \tag{6-2-22}$$

The magnitude response $20 \log |H(\omega)|$ as a function of the normalized frequency ω/ω_n is illustrated in Fig. 6-2-5, with the damping factor ζ as a parameter and $\tau_1 \gg 1$. Note that $\zeta = 1$ results in a critically damped loop response, $\zeta < 1$ produces an underdamped response, and $\zeta > 1$ yields an overdamped response.

In practice, the selection of the bandwidth of the PLL involves a trade-off between speed of response and noise in the phase estimate, which is the topic considered below. On the one hand, it is desirable to select the bandwidth of the loop to be sufficiently wide to track any time variations in the phase of the received carrier. On the other, a wideband PLL allows more noise to pass into the loop, which corrupts the phase estimate. Below, we assess the effects of noise in the quality of the phase estimate.

6-2-3 Effect of Additive Noise on the Phase Estimate

In order to evaluate the effects of noise on the estimate of the carrier phase, let us assume that the noise at the input to the PLL is narrowband. For this analysis, we assume that the PLL is tracking a sinusoidal signal of the form

$$s(t) = A_c \cos [2\pi f_c t + \phi(t)] \tag{6-2-23}$$

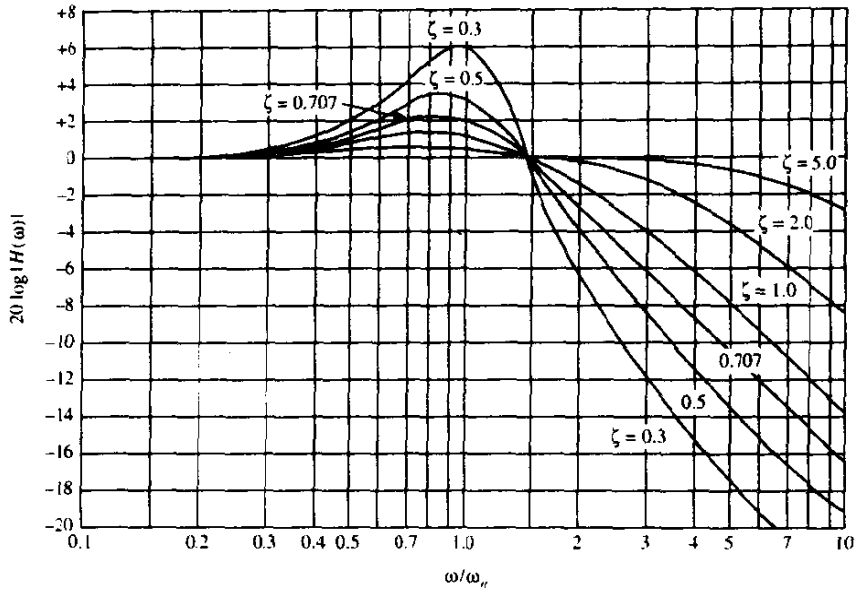


FIGURE 6-2-5 Frequency response of a second-order loop. [*From* Phaselock Techniques, *2nd edition, by F. M. Gardner, © 1979 by John Wiley and Sons, Inc. Reprinted with permission of the publisher.*]

that is corrupted by the additive narrowband noise

$$n(t) = x(t)\cos 2\pi f_c t - y(t)\sin 2\pi f_c t \tag{6-2-24}$$

The in-phase and quadrature components of the noise are assumed to be statistically independent, stationary gaussian noise processes with (two-sided) power spectral density $\frac{1}{2}N_0$ W/Hz. By using simple trigonometric identities, the noise term in (6-2-24) can be expressed as

$$n(t) = n_c(t)\cos[2\pi f_c t + \phi(t)] - n_s(t)\sin[2\pi f_c t + \phi(t)] \tag{6-2-25}$$

where

$$\begin{aligned} n_c(t) &= x(t)\cos\phi(t) + y(t)\sin\phi(t) \\ n_s(t) &= -x(t)\sin\phi(t) + y(t)\cos\phi(t) \end{aligned} \tag{6-2-26}$$

We note that

$$n_c(t) + jn_s(t) = [x(t) + jy(t)]e^{-j\phi(t)}$$

so that the quadrature components $n_c(t)$ and $n_s(t)$ have exactly the same statistical characteristics as $x(t)$ and $y(t)$.

If $s(t) + n(t)$ is multiplied by the output of the VCO and the double-frequency terms are neglected, the input to the loop filter is the noise-corrupted signal

$$\begin{aligned} e(t) &= A_c\sin\Delta\phi + n_c(t)\sin\Delta\phi - n_s(t)\cos\Delta\phi \\ &= A_c\sin\Delta\phi + n_1(t) \end{aligned} \tag{6-2-27}$$

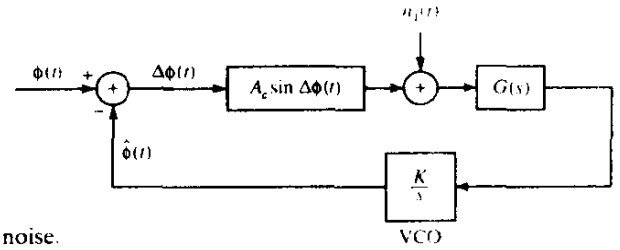


FIGURE 6-2-6 Equivalent PLL model with additive noise.

where, by definition, $\Delta\phi = \phi - \hat{\phi}$ is the phase error. Thus, we have the equivalent model for the PLL with additive noise as shown in Fig. 6-2-6.

When the power $P_c = \frac{1}{2}A_c^2$ of the incoming signal is much larger than the noise power, we may linearize the PLL and, thus, easily determine the effect of the additive noise on the quality of the estimate $\hat{\phi}$. Under these conditions, the model for the linearized PLL with additive noise is illustrated in Fig. 6-2-7. Note that the gain parameter A_c may be normalized to unity, provided that the noise terms are scaled by $1/A_c$, i.e., the noise terms become

$$n_2(t) = \frac{n_c(t)}{A_c} \sin \Delta\phi - \frac{n_s(t)}{A_c} \cos \Delta\phi \tag{6-2-28}$$

Since the noise $n_2(t)$ is additive at the input to the loop, the variance of the phase error $\Delta\phi$, which is also the variance of the VCO output phase, is

$$\sigma_{\hat{\phi}}^2 = \frac{N_0 B_{eq}}{A_c^2} \tag{6-2-29}$$

where B_{eq} is the (one-sided) equivalent noise bandwidth of the loop, given in (6-2-22). Note that $\sigma_{\hat{\phi}}^2$ is simply the ratio of total noise power within the bandwidth of the PLL divided by the signal power A^2. Hence,

$$\sigma_{\hat{\phi}}^2 = 1/\gamma_L \tag{6-2-30}$$

where γ_L is defined as the signal-to-noise ratio

$$\text{SNR} \equiv \gamma_L = \frac{A_c^2}{N_0 B_{eq}} \tag{6-2-31}$$

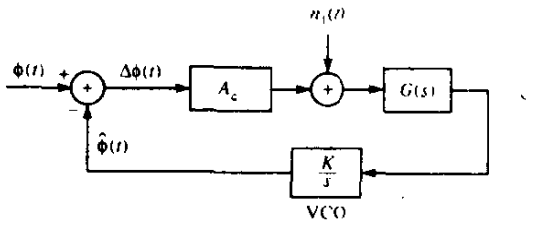


FIGURE 6-2-7 Linearized PLL model with additive noise.

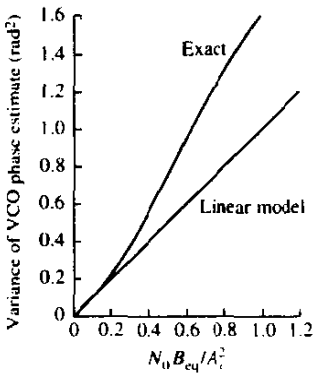


FIGURE 6-2-8 Comparison of VCO phase variance for exact and approximate (linear model) first-order PLL. [*From* Principles of Coherent Communication, *by A. J. Viterbi; © 1966 by McGraw-Hill Book Company. Reprinted with permission of the publisher.*]

The expression for the variance σ_ϕ^2 of the VCO phase error applies to the case where the SNR is sufficiently high that the linear model for the PLL applies. An exact analysis based on the nonlinear PLL is mathematically tractable when $G(s) = 1$, which results in a first-order loop. In this case, the probability density function for the phase error may be derived (see Viterbi, 1966) and has the form

$$p(\Delta\phi) = \frac{\exp(\gamma_L \cos \Delta\phi)}{2\pi I_0(\gamma_L)} \tag{6-2-32}$$

where γ_L is the SNR given by (6-2-31) with B_{eq} being the appropriate noise bandwidth of the first-order loop, and $I_0(\cdot)$ is the modified Bessel function of order zero.

From the expression for $p(\Delta\phi)$, we may obtain the exact value of the variance for the phase error on a first-order PLL. This is plotted in Fig. 6-2-8 as a function of $1/\gamma_L$. Also shown for comparison is the result obtained with the linearized PLL model. Note that the variance for the linear model is close to the exact variance for $\gamma_L > 3$. Hence, the linear model is adequate for practical purposes.

Approximate analyses of the statistical characteristics of the phase error for the nonlinear PLL have also been performed. Of particular importance is the transient behavior of the PLL during initial acquisition. Another important problem is the behavior of PLL at low SNR. It is known, for example, that when the SNR at the input to the PLL drops below a certain value, there is a rapid deterioration in the performance of the PLL. The loop begins to lose lock and an impulsive-type of noise, characterized as clicks, is generated which degrades the performance of the loop. Results on these topics can be found in the texts by Viterbi (1966), Lindsey (1972), Lindsey and Simon (1973), and Gardner (1979), and in the survey papers by Gupta (1975) and Lindsey and Chie (1981).

Up to this point, we have considered carrier phase estimation when the carrier signal is unmodulated. Below, we consider carrier phase recovery when the signal carries information.

6-2-4 Decision-Directed Loops

A problem arises in maximizing either (6-2-9) or (6-2-10) when the signal $s(t; \phi)$ carries the information sequence $\{I_n\}$. In this case we can adopt one of two approaches: either we assume that $\{I_n\}$ is known or we treat $\{I_n\}$ as a random sequence and average over its statistics.

In decision-directed parameter estimation, we assume that the information sequence $\{I_n\}$ over the observation interval has been estimated and, in the absence of demodulation errors, $\tilde{I}_n = I_n$, where $\tilde{I}_n$ denotes the detected value of the information I_n. In this case $s(t; \phi)$ is completely known except for the carrier phase. Decision-directed phase estimation was first described by Proakis *et al.* (1964).

To be specific, let us consider the decision-directed phase estimate for the class of linear modulation techniques for which the received *equivalent lowpass signal* may be expressed as

$$\begin{aligned} r(t) &= e^{-j\phi} \sum_n I_n g(t - nT) + z(t) \\ &= s_l(t) e^{-j\phi} + z(t) \end{aligned} \tag{6-2-33}$$

where $s_l(t)$ is a known signal if the sequence $\{I_n\}$ is assumed known. The likelihood function and corresponding log-likelihood function for the equivalent lowpass signal are

$$\Lambda(\phi) = C \exp\left\{\text{Re}\left[\frac{1}{N_0}\int_{T_0} r(t) s_l^*(t) e^{j\phi}\, dt\right]\right\} \tag{6-2-34}$$

$$\Lambda_L(\phi) = \text{Re}\left\{\left[\frac{1}{N_0}\int_{T_0} r(t) s_l^*(t)\, dt\right] e^{j\phi}\right\} \tag{6-2-35}$$

If we substitute for $s_l(t)$ in (6-2-35) and assume that the observation interval $T_0 = KT$, where K is a positive integer, we obtain

$$\begin{aligned} \Lambda_L(\phi) &= \text{Re}\left\{e^{j\phi} \frac{1}{N_0} \sum_{n=0}^{K-1} I_n^* \int_{nT}^{(n+1)T} r(t) g^*(t - nT)\, dt\right\} \\ &= \text{Re}\left\{e^{j\phi} \frac{1}{N_0} \sum_{n=0}^{K-1} I_n^* y_n\right\} \end{aligned} \tag{6-2-36}$$

where, by definition

$$y_n = \int_{nT}^{(n+1)T} r(t) g^*(t - nT)\, dt \tag{6-2-37}$$

Note that y_n is the output of the matched filter in the nth signal interval. The ML estimate of ϕ is easily found from (6-2-36) by differentiating the log-likelihood

$$\Lambda_L(\phi) = \text{Re}\left(\frac{1}{N_0} \sum_{n=0}^{K-1} I_n^* y_n\right) \cos\phi - \text{Im}\left(\frac{1}{N_0} \sum_{n=0}^{K-1} I_n^* y_n\right) \sin\phi$$

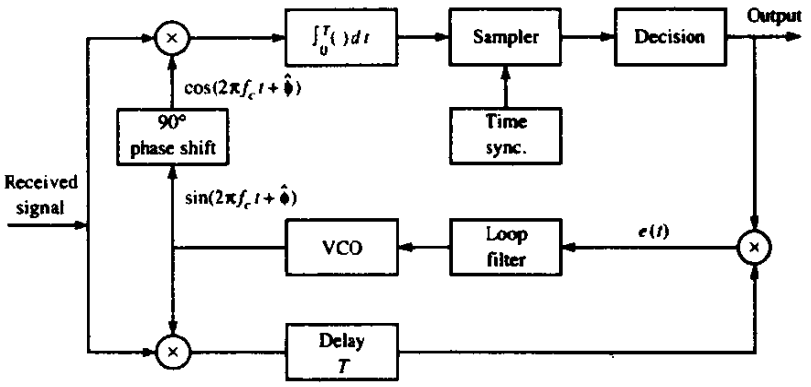


FIGURE 6-2-9 Carrier recovery with a decision-feedback PLL.

with respect to ϕ and setting the derivative equal to zero. Thus, we obtain

$$\hat{\phi}_{\mathrm{ML}} = -\tan^{-1}\left[\mathrm{Im}\left(\sum_{n=0}^{K-1} I_n^* y_n\right) \Big/ \mathrm{Re}\left(\sum_{n=0}^{K-1} I_n^* y_n\right)\right] \tag{6-2-38}$$

We call $\hat{\phi}_{\mathrm{ML}}$ in (6-2-38) the *decision-directed* (or *decision-feedback*) *carrier phase estimate.* It is easily shown (Problem 6-10) that the mean value of $\hat{\phi}_{\mathrm{ML}}$ is ϕ, so that the estimate is unbiased. Furthermore, the pdf of $\hat{\phi}_{\mathrm{ML}}$ can be obtained (Problem 6-11) by using the procedure described in Section 5-2-7.

A decision-feedback PLL (DFPLL) that is appropriate for a double-sideband PAM signal of the form $A(t)\cos(2\pi f_c t + \phi)$ is shown in Fig. 6-2-9. The received signal is multiplied by the quadrature carriers $c_c(t)$ and $c_s(t)$, as given by (6-2-5), which are derived from the VCO. The product signal

$$\begin{aligned} r(t)\cos(2\pi f_c t + \hat{\phi}) = {} & \tfrac{1}{2}[A(t) + n_c(t)]\cos\Delta\phi \\ & - \tfrac{1}{2}n_s(t)\sin\Delta\phi + \text{double-frequency terms} \end{aligned} \tag{6-2-39}$$

is used to recover the information carried by $A(t)$. The detector makes a decision on the symbol that is received every T seconds. Thus, in the absence of decision errors, it reconstructs $A(t)$ free of any noise. This reconstructed signal is used to multiply the product of the second quadrature multiplier, which has been delayed by T seconds to allow the demodulator to reach a decision. Thus, the input to the loop filter in the absence of decision errors is the error signal

$$\begin{aligned} e(t) &= \tfrac{1}{2}A(t)\{[A(t) + n_c(t)]\sin\Delta\phi - n_s(t)\cos\Delta\phi\} \\ &\quad + \text{double-frequency terms} \\ &= \tfrac{1}{2}A^2(t)\sin\Delta\phi + \tfrac{1}{2}A(t)[n_c(t)\sin\Delta\phi - n_s(t)\cos\Delta\phi] \\ &\quad + \text{double-frequency terms} \end{aligned} \tag{6-2-40}$$

The loop filter is lowpass and, hence, it rejects the double-frequency term in $e(t)$. The desired component is $A^2(t)\sin\Delta\phi$, which contains the phase error for driving the loop.

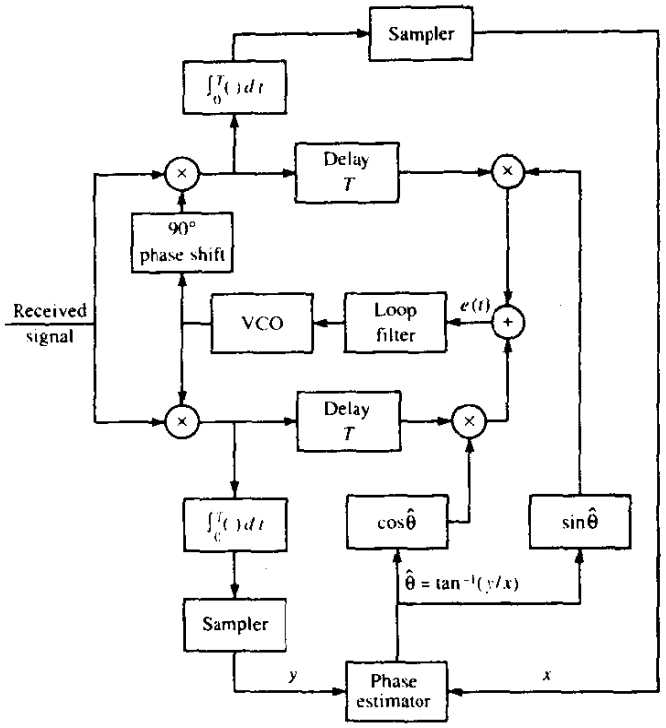


FIGURE 6-2-10 Carrier recovery for M-ary PSK using a decision-feedback PLL.

In the case of M-ary PSK, the DFPLL has the configuration shown in Fig. 6-2-10. The received signal is demodulated to yield the phase estimate

$$\hat{\theta}_m = \frac{2\pi}{M}(m-1)$$

which, in the absence of a decision error, is the transmitted signal phase. The two outputs of the quadrature multipliers are delayed by the symbol duration T and multiplied by $\cos\theta_m$ and $\sin\theta_m$ to yield

$$\begin{aligned} &r(t)\cos(2\pi f_c t + \hat{\phi})\sin\theta_m \\ &\quad = \tfrac{1}{2}[A\cos\theta_m + n_c(t)]\sin\theta_m\cos(\phi-\hat{\phi}) \\ &\qquad - \tfrac{1}{2}[A\sin\theta_m + n_s(t)]\sin\theta_m\sin(\phi-\hat{\phi}) \\ &\qquad + \text{double-frequency terms} \\ &r(t)\sin(2\pi f_c t + \hat{\phi})\cos\theta_m \\ &\quad = -\tfrac{1}{2}[A\cos\theta_m + n_c(t)]\cos\theta_m\sin(\phi-\hat{\phi}) \\ &\qquad - \tfrac{1}{2}[A\sin\theta_m + n_s(t)]\cos\theta_m\cos(\phi-\hat{\phi}) \\ &\qquad + \text{double-frequence terms} \end{aligned} \tag{6-2-41}$$

The two signals are added to generate the error signal

$$e(t) = -\tfrac{1}{2}A \sin(\phi - \hat{\phi}) + \tfrac{1}{2}n_c(t) \sin(\phi - \hat{\phi} - \theta_m) + \tfrac{1}{2}n_s(t) \cos(\phi - \hat{\phi} - \theta_m) + \text{double-frequency terms} \quad (6\text{-}2\text{-}42)$$

This error signal is the input to the loop filter that provides the control signal for the VCO.

We observe that the two quadrature noise components in (6-2-42) appear as additive terms. There is no term involving a product of two noise components as in an Mth-power law device, described in the next section. Consequently, there is no additional power loss associated with the decision-feedback PLL.

This M-phase tracking loop has a phase ambiguity of $360°/M$, necessitating the need to differentially encode the information sequence prior to transmission and differentially decode the received sequence after demodulation to recover the information.

The ML estimate in (6-2-38) is also appropriate for QAM. The ML estimate for offset QPSK is also easily obtained (Problem 6-12) by maximizing the log-likelihood function in (6-2-35), with $s_l(t)$ given as

$$s_l(t) = \sum_n I_n g(t - nT) + j \sum_n J_n g(t - nT - \tfrac{1}{2}T) \quad (6\text{-}2\text{-}43)$$

where $I_n = \pm 1$ and $J_n = \pm 1$.

Finally, we should also mention that carrier phase recovery for CPM signals can be accomplished in a decision-directed manner by use of a PLL. From the optimum demodulator for CPM signals, which is described in Section 5-3, we can generate an error signal that is filtered in a loop filter whose output drives a PLL.

6-2-5 Non-Decision-Directed Loops

Instead of using a decision-directed scheme to obtain the phase estimate, we may treat the data as random variables and simply average $\Lambda(\phi)$ over these random variables prior to maximization. In order to carry out this integration, we may use either the actual probability distribution function of the data, if it is known or, perhaps, we may assume some probability distribution that might be a reasonable approximation to the true distribution. The following example illustrates the first approach.

Example 6-2-2

Suppose the real signal $s(t)$ carries binary modulation. Then, in a signal interval, we have

$$s(t) = A \cos 2\pi f_c t, \quad 0 \leq t \leq T$$

where $A = \pm 1$ with equal probability. Clearly, the pdf of A is given as

$$p(A) = \tfrac{1}{2}\delta(A-1) + \tfrac{1}{2}\delta(A+1)$$

Now, the likelihood function $\Lambda(\phi)$ given by (6-2-9) is conditional on a given value of A and must be averaged over the two values. Thus,

$$\begin{aligned}\bar{\Lambda}(\phi) &= \int_{-\infty}^{\infty} \Lambda(\phi) p(A)\, dA \\ &= \tfrac{1}{2}\exp\left[\frac{2}{N_0}\int_0^T r(t)\cos(2\pi f_c t + \phi)\, dt\right] \\ &\quad + \tfrac{1}{2}\exp\left[-\frac{2}{N_0}\int_0^T r(t)\cos(2\pi f_c t + \phi)\, dt\right] \\ &= \cosh\left[\frac{2}{N_0}\int_0^T r(t)\cos(2\pi f_c t + \phi)\, dt\right]\end{aligned}$$

and the corresponding log-likelihood function is

$$\bar{\Lambda}_L(\phi) = \ln\cosh\left[\frac{2}{N_0}\int_0^T r(t)\cos(2\pi f_c t + \phi)\, dt\right] \tag{6-2-44}$$

If we differentiate $\bar{\Lambda}_L(\phi)$ and set the derivative equal to zero, we obtain the ML estimate for the non-decision-directed estimate. Unfortunately, the functional relationship in (6-2-44) is highly nonlinear and, hence, an exact solution is difficult to obtain. On the other hand, approximations are possible. In particular,

$$\ln\cosh x = \begin{cases} \frac{1}{2}x^2 & (|x| \ll 1) \\ |x| & (|x| \gg 1) \end{cases} \tag{6-2-45}$$

With these approximations, the solution for ϕ becomes tractable.

In this example, we averaged over the two possible values of the information symbol. When the information symbols are M-valued, where M is large, the averaging operation yields highly nonlinear functions of the parameter to be estimated. In such a case, we may simplify the problem by assuming that the information symbols are continuous random variables. For example, we may assume that the symbols are zero-mean gaussian. The following example illustrates this approximation and the resulting form for the average likelihood function.

Example 6-2-3

Let us consider the same signal as in Example 6-2-2, but now we assume that the amplitude A is zero-mean gaussian with unit variance. Thus,

$$p(A) = \frac{1}{\sqrt{2\pi}} e^{-A^2/2}$$

If we average $\Lambda(\phi)$ over the assumed pdf of A, we obtain the average likelihood $\bar{\Lambda}(\phi)$ in the form

$$\bar{\Lambda}(\phi) = C \exp\left\{\left[\frac{2}{N_0}\int_0^T r(t)\cos(2\pi f_c t + \phi)\,dt\right]^2\right\} \tag{6-2-46}$$

and the corresponding log-likelihood as

$$\bar{\Lambda}_L(\phi) = \left[\frac{2}{N_0}\int_0^T r(t)\cos(2\pi f_c t + \phi)\,dt\right]^2 \tag{6-2-47}$$

We can obtain the ML estimate of ϕ by differentiating $\bar{\Lambda}_L(\phi)$ and setting the derivative to zero.

It is interesting to note that the log-likelihood function is quadratic under the gaussian assumption and that it is approximately quadratic, as indicated in (6-2-45) for small values of the cross-correlation of $r(t)$ with $s(t;\phi)$. In other words, if the cross-correlation over a single interval is small, the gaussian assumption for the distribution of the information symbols yields a good approximation to the log-likelihood function.

In view of these results, we may use the gaussian approximation on all the symbols in the observation interval $T_0 = KT$. Specifically, we assume that the K information symbols are statistically independent and identically distributed. By averaging the likelihood function $\Lambda(\phi)$ over the gaussian pdf for each of the K symbols in the interval $T_0 = KT$, we obtain the result

$$\Lambda(\phi) = C \exp\left\{\sum_{n=0}^{K-1}\left[\frac{2}{N_0}\int_{nT}^{(n+1)T} r(t)\cos(2\pi f_c t + \phi)\,dt\right]^2\right\} \tag{6-2-48}$$

If we take the logarithm of (6-2-48), differentiate the resulting log-likelihood function, and set the derivative equal to zero, we obtain the condition for the ML estimate as

$$\sum_{n=0}^{K-1}\int_{nT}^{(n+1)T} r(t)\cos(2\pi f_c t + \hat{\phi})\,dt \int_{nT}^{(n+1)T} r(t)\sin(2\pi f_c t + \hat{\phi})\,dt = 0 \tag{6-2-49}$$

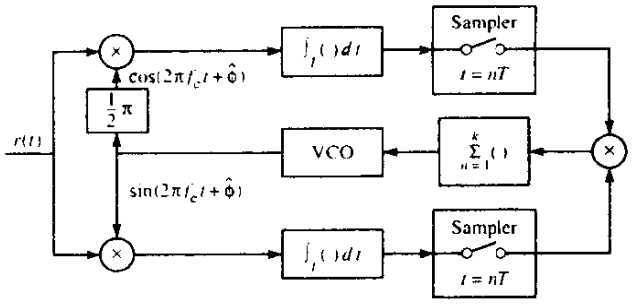


FIGURE 6-2-11 Non-decision-directed PLL for carrier phase estimations of PAM signals.

Although this equation can be manipulated further, its present form suggests the tracking loop configuration illustrated in Fig. 6-2-11. This loop resembles a Costas loop, which is described below. We note that the multiplication of the two signals from the integrators destroys the sign carried by the information symbols. The summer plays the role of the loop filter. In a tracking loop configuration, the summer may be implemented either as a sliding-window digital filter (summer) or as a lowpass digital filter with exponential weighting of the past data.

In a similar manner, one can derive non-decision directed ML phase estimates for QAM and M-PSK. The starting point is to average the likelihood function given by (6-2-9) over the statistical characteristics of the data. Here again, we may use the gaussian approximation (two-dimensional gaussian for complex-valued information symbols) in averaging over the information sequence.

Squaring Loop The squaring loop is a non-decision-directed loop that is widely used in practice to establish the carrier phase of double-sideband suppressed carrier signals such as PAM. To describe its operation, consider the problem of estimating the carrier phase of the digitally modulated PAM signal of the form

$$s(t) = A(t)\cos(2\pi f_c t + \phi) \tag{6-2-50}$$

where $A(t)$ carries the digital information. Note that $E[s(t)] = E[A(t)] = 0$ when the signal levels are symmetric about zero. Consequently, the average value of $s(t)$ does not produce any phase coherent frequency components at any frequency, including the carrier. One method for generating a carrier from the received signal is to square the signal and, thus, to generate a frequency component at $2f_c$, which can be used to drive a phase-locked loop (PLL) tuned to $2f_c$. This method is illustrated in the block diagram shown in Fig. 6-2-12.

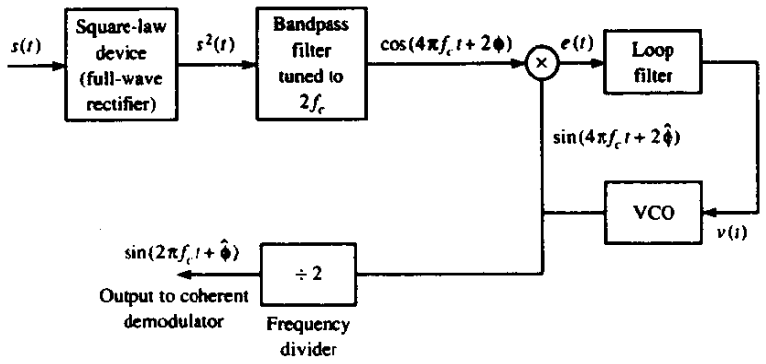


FIGURE 6-2-12 Carrier recovery using a square-law device.

The output of the square-law device is

$$s^2(t) = A^2(t)\cos^2(2\pi f_c t + \phi)$$
$$= \tfrac{1}{2}A^2(t) + \tfrac{1}{2}A^2(t)\cos(4\pi f_c t + 2\phi) \qquad (6\text{-}2\text{-}51)$$

Since the modulation is a cyclostationary stochastic process, the expected value of $s^2(t)$ is

$$E[s^2(t)] = \tfrac{1}{2}E[A^2(t)] + \tfrac{1}{2}E[A^2(t)]\cos(4\pi f_c t + 2\phi) \qquad (6\text{-}2\text{-}52)$$

Hence, there is power at the frequency $2f_c$.

If the output of the square-law device is passed through a bandpass filter tuned to the double-frequency term in (6-2-51), the mean value of the filter is a sinusoid with frequency $2f_c$, phase 2ϕ, and amplitude $\frac{1}{2}E[A^2(t)]H(2f_c)$, where $H(2f_c)$ is the gain of the filter at $f = 2f_c$. Thus, the square-law device has produced a periodic component from the input signal $s(t)$. In effect, the squaring of $s(t)$ has removed the sign information contained in $A(t)$ and, thus, has resulted in phase-coherent frequency components at twice the carrier. The filtered frequency component at $2f_c$ is then used to drive the PLL.

The squaring operation leads to a noise enhancement that increases the noise power level at the input to the PLL and results in an increase in the variance of the phase error.

To elaborate on this point, let the input to the squarer be $s(t) + n(t)$, where $s(t)$ is given by (6-2-50) and $n(t)$ represents the bandpass additive gaussian noise process. By squaring $s(t) + n(t)$, we obtain

$$y(t) = s^2(t) + 2s(t)n(t) + n^2(t) \qquad (6\text{-}2\text{-}53)$$

where $s^2(t)$ is the desired signal component and the other two components are the signal × noise and noise × noise terms. By computing the autocorrelation functions and power density spectra of these two noise components, one can

easily show that both components have spectral power in the frequency band centered at $2f_c$. Consequently, the bandpass filter with bandwidth B_{bp} centered at $2f_c$, which produces the desired sinusoidal signal component that drives the PLL, also passes noise due to these two terms.

Since the bandwidth of the loop is designed to be significantly smaller than the bandwidth B_{bp} of the bandpass filter, the total noise spectrum at the input to the PLL may be approximated as a constant within the loop bandwidth. This approximation allows us to obtain a simple expression for the variance of the phase error as

$$\sigma_{\hat{\phi}}^2 = 1/\gamma_L S_L \tag{6-2-54}$$

where S_L is called the squaring loss and is given by

$$S_L = \left(1 + \frac{B_{bp}/2B_{eq}}{\gamma_L}\right)^{-1} \tag{6-2-55}$$

Since $S_L < 1$, S_L^{-1} represents the increase in the variance of the phase error caused by the added noise (noise × noise terms) that results from the squarer. Note, for example, that when $\gamma_L = B_{bp}/2B_{eq}$, the loss is 3 dB.

Finally, we observe that the output of the VCO from the squaring loop must be frequency-divided by 2 to generate the phase-locked carrier for signal demodulation. It should be noted that the output of the frequency divider has a phase ambiguity of 180° relative to the phase of the received signal. For this reason, the binary data must be differentially encoded prior to transmission and differentially decoded at the receiver.

Costas Loop Another method for generating a properly phased carrier for a double-sideband suppressed carrier signal is illustrated by the block diagram shown in Fig. 6-2-13. This scheme was developed by Costas (1956) and is

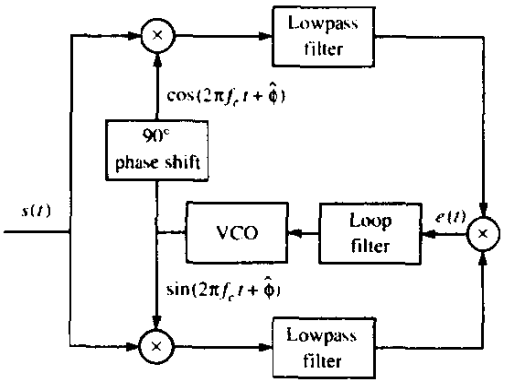


FIGURE 6-2-13 Block diagram of Costas loop.

called the *Costas loop*. The received signal is multiplied by $\cos(2\pi f_c t + \hat{\phi})$ and $\sin(2\pi f_c t + \hat{\phi})$, which are outputs from the VCO. The two products are

$$\begin{aligned} y_c(t) &= [s(t) + n(t)] \cos(2\pi f_c t + \hat{\phi}) \\ &= \tfrac{1}{2}[A(t) + n_c(t)] \cos \Delta\phi + \tfrac{1}{2} n_s(t) \sin \Delta\phi \\ &\quad + \text{double-frequency terms} \\ y_s(t) &= [s(t) + n(t)] \sin(2\pi f_c t + \hat{\phi}) \\ &= \tfrac{1}{2}[A(t) + n_c(t)] \sin \Delta\phi - \tfrac{1}{2} n_s(t) \cos \Delta\phi \\ &\quad + \text{double-frequency terms} \end{aligned} \tag{6-2-56}$$

where the phase error $\Delta\phi = \hat{\phi} - \phi$. The double-frequency terms are eliminated by the lowpass filters following the multiplications.

An error signal is generated by multiplying the two outputs of the lowpass filters. Thus,

$$\begin{aligned} e(t) &= \tfrac{1}{8}\{[A(t) + n_c(t)]^2 - n_s^2(t)\} \sin(2\,\Delta\phi) \\ &\quad - \tfrac{1}{4} n_s(t)[A(t) + n_c(t)] \cos(2\,\Delta\phi) \end{aligned} \tag{6-2-57}$$

This error signal is filtered by the loop filter, whose output is the control voltage that drives the VCO. The reader should note the similarity of the Costas loop to the PLL shown in Fig. 6-2-11.

We note that the error signal into the loop filter consists of the desired term $A^2(t) \sin 2(\hat{\phi} - \phi)$ plus terms that involve signal × noise and noise × noise. These terms are similar to the two noise terms at the input to the PLL for the squaring method. In fact, if the loop filter in the Costas loop is identical to that used in the squaring loop, the two loops are equivalent. Under this condition, the probability density function of the phase error and the performance of the two loops are identical.

It is interesting to note that the optimum lowpass filter for rejecting the double-frequency terms in the Costas loop is a filter matched to the signal pulse in the information-bearing signal. If matched filters are employed for the low pass filters, their outputs could be sampled at the bit rate, at the end of each signal interval, and the discrete-time signal samples could be used to drive the loop. The use of the matched filter results in a smaller noise into the loop.

Finally, we note that, as in the squaring PLL, the output of the VCO contains a phase ambiguity of 180°, necessitating the need for differential encoding of the data prior to transmission and differential decoding at the demodulator.

Carrier Estimation for Multiple Phase Signals When the digital information is transmitted via M-phase modulation of a carrier, the methods described above can be generalized to provide the properly phased carrier for

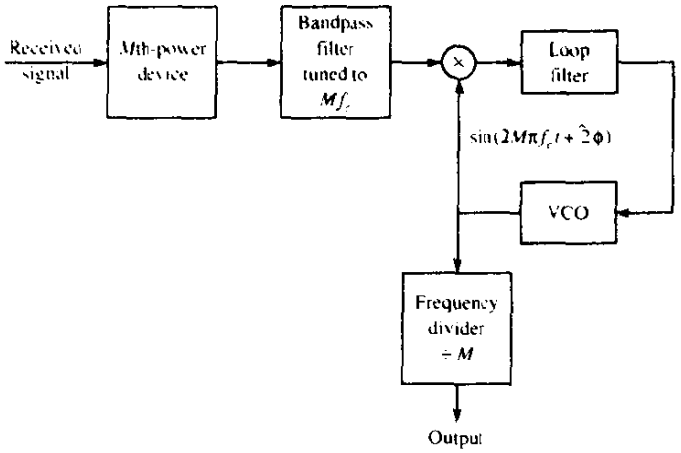


FIGURE 6-2-14 *Carrier recovery with Mth power law device for M-ary PSK.*

demodulation. The received M-phase signal, excluding the additive noise, may be expressed as

$$s(t) = A\cos\left[2\pi f_c t + \phi + \frac{2\pi}{M}(m-1)\right], \quad m = 1, 2, \ldots, M \qquad (6\text{-}2\text{-}58)$$

where $2\pi(m-1)/M$ represents the information-bearing component of the signal phase. The problem in carrier recovery is to remove the information-bearing component and, thus, to obtain the unmodulated carrier $\cos(2\pi f_c t + \phi)$. One method by which this can be accomplished is illustrated in Fig. 6-2-14, which represents a generalization of the squaring loop. The signal is passed through an Mth-power-law device, which generates a number of harmonics of f_c. The bandpass filter selects the harmonic $\cos(2\pi M f_c t + M\phi)$ for driving the PLL. The term

$$\frac{2\pi}{M}(m-1)M = 2\pi(m-1) \equiv 0 \pmod{2\pi}, \quad m = 1, 2, \ldots, M$$

Thus, the information is removed. The VCO output is $\sin(2\pi M f_c t + M\hat{\phi})$, so this output is divided in frequency by M to yield $\sin(2\pi f_c t + \hat{\phi})$, and phase-shifted by $\frac{1}{2}\pi$ rad to yield $\cos(2\pi f_c t + \hat{\phi})$. These components are then fed to the demodulator. Although not explicitly shown, there is a phase ambiguity in these reference sinusoids of $360°/M$, which can be overcome by differential encoding of the data at the transmitter and differential decoding after demodulation at the receiver.

Just as in the case of the squaring PLL, the Mth-power PLL operates in the presence of noise that has been enhanced by the Mth-power-law device, which results in the output

$$y(t) = [s(t) + n(t)]^M$$

The variance of the phase error in the PLL resulting from the additive noise may be expressed in the simple form

$$\sigma_{\hat{\phi}}^2 = \frac{S_{\text{ML}}^{-1}}{\gamma_L} \tag{6-2-59}$$

where γ_L is the loop SNR and S_{ML}^{-1} is the *M-phase power loss*. S_{ML} has been evaluated by Lindsey and Simon (1973) for $M = 4$ and 8.

Another method for carrier recovery in M-ary PSK is based on a generalization of the Costas loop. That method requires multiplying the received signal by M phase-shifted carriers of the form

$$\sin\left[2\pi f_c t + \hat{\phi} + \frac{\pi}{M}(k-1)\right], \quad k = 1, 2, \ldots, M$$

lowpass-filtering each product, and then multiplying the outputs of the lowpass filters to generate the error signal. The error signal excites the loop filter, which, in turn, provides the control signal for the VCO. This method is relatively complex to implement and, consequently, has not been generally used in practice.

Comparison of Decision-Directed with Non-Decision-Directed Loops We note that the decision-feedback phase-locked loop (DFPLL) differs from the Costas loop only in the method by which $A(t)$ is rectified for the purpose of removing the modulation. In the Costas loop, each of the two quadrature signals used to rectify $A(t)$ is corrupted by noise. In the DFPLL, only one of the signals used to rectify $A(t)$ is corrupted by noise. On the other hand, the squaring loop is similar to the Costas loop in terms of the noise effect on the estimate $\hat{\phi}$. Consequently, the DFPLL is superior in performance to both the Costas loop and the squaring loop, provided that the demodulator is operating at error rates below 10^{-2} where an occasional decision error has a negligible effect on $\hat{\phi}$. Quantitative comparisons of the variance of the phase errors in a Costas loop to those in a DFPLL have been made by Lindsey and Simon (1973), and show that the variance of the DFPLL is 4–10 times smaller for signal-to-noise ratios per bit above 0 db.

6-3 SYMBOL TIMING ESTIMATION

In a digital communication system, the output of the demodulator must be sampled periodically at the symbol rate, at the precise sampling time instants $t_m = mT + \tau$, where T is the symbol interval and τ is a nominal time delay that accounts for the propagation time of the signal from the transmitter to the receiver. To perform this periodic sampling, we require a clock signal at the

receiver. The process of extracting such a clock signal at the receiver is usually called *symbol synchronization* or *timing recovery*.

Timing recovery is one of the most critical functions that is performed at the receiver of a synchronous digital communication system. We should note that the receiver must know not only the frequency $(1/T)$ at which the outputs of the matched filters or correlators are sampled, but also where to take the samples within each symbol interval. The choice of sampling instant within the symbol interval of duration T is called the *timing phase*.

Symbol synchronization can be accomplished in one of several ways. In some communication systems, the transmitter and receiver clocks are synchronized to a master clock, which provides a very precise timing signal. In this case, the receiver must estimate and compensate for the relative time delay between the transmitted and received signals. Such may be the case for radio communication systems that operate in the very low frequency (VLF) band (below 30 kHz), where precise clock signals are transmitted from a master radio station.

Another method for achieving symbol synchronization is for the transmitter to simultaneously transmit the clock frequency $1/T$ or a multiple of $1/T$ along with the information signal. The receiver may simply employ a narrowband filter tuned to the transmitted clock frequency and, thus, extract the clock signal for sampling. This approach has the advantage of being simple to implement. There are several disadvantages, however. One is that the transmitter must allocate some of its available power to the transmission of the clock signal. Another is that some small fraction of the available channel bandwidth must be allocated for the transmission of the clock signal. In spite of these disadvantages, this method is frequently used in telephone transmission systems that employ large bandwidths to transmit the signals of many users. In such a case, the transmission of a clock signal is shared in the demodulation of the signals among the many users. Through this shared use of the clock signal, the penalty in transmitter power and in bandwidth allocation is reduced proportionally by the number of users.

A clock signal can also be extracted from the received data signal. There are a number of different methods that can be used at the receiver to achieve self-synchronization. In this section, we treat both decision-directed and non-decision-directed methods.

6-3-1 Maximum-Likelihood Timing Estimation

Let us begin by obtaining the ML estimate of the time delay τ. If the signal is a baseband PAM waveform, it is represented as

$$r(t) = s(t; \tau) + n(t) \tag{6-3-1}$$

where

$$s(t; \tau) = \sum_n I_n g(t - nT - \tau) \tag{6-3-2}$$

As in the case of ML phase estimation, we distinguish between two types of timing estimators, decision-directed timing estimators and non-decision-directed estimators. In the former, the information symbols from the output of the demodulator are treated as the known transmitted sequence. In this case, the log-likelihood function has the form

$$\Lambda_L(\tau) = C_L \int_{T_0} r(t)s(t;\tau)\,dt \tag{6-3-3}$$

If we substitute (6-3-2) into (6-3-3), we obtain

$$\begin{aligned}\Lambda_L(\tau) &= C_L \sum_n I_n \int_{T_0} r(t)g(t-nT-\tau)\,dt \\ &= C_L \sum_n I_n y_n(\tau)\end{aligned} \tag{6-3-4}$$

where $y_n(t)$ is defined as

$$y_n(\tau) = \int_{T_0} r(t)g(t-nT-\tau)\,dt \tag{6-3-5}$$

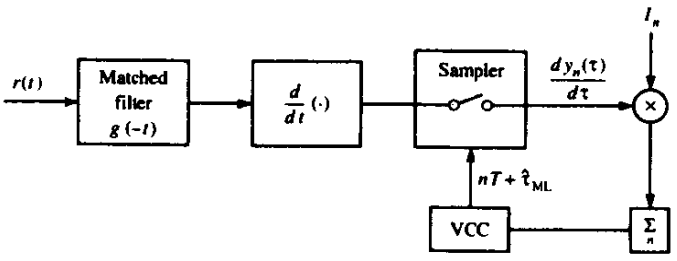
A necessary condition for $\hat{\tau}$ to be the ML estimate of τ is that

$$\begin{aligned}\frac{d\Lambda_L(\tau)}{d\tau} &= \sum_n I_n \frac{d}{d\tau}\int_{T_0} r(t)g(t-nT-\tau)\,dt \\ &= \sum_n I_n \frac{d}{d\tau}[y_n(\tau)] = 0\end{aligned} \tag{6-3-6}$$

The result in (6-3-6) suggests the implementation of the tracking loop shown in Fig. 6-3-1. We should observe that the summation in the loop serves as the loop filter whose bandwidth is controlled by the length of the sliding window in the summation. The output of the loop filter drives the voltage-controlled clock (VCC), or voltage-controlled oscillator, which controls the sampling times for the input to the loop. Since the detected information sequence $\{I_n\}$ is used in the estimation of τ, the estimate is decision-directed.

The techniques described above for ML timing estimation of baseband

FIGURE 6-3-1 Decision-directed ML estimation of timing for baseband PAM.



PAM signals can be extended to carrier modulated signal formats such as QAM and PSK in a straightforward manner, by dealing with the equivalent lowpass form of the signals. Thus, the problem of ML estimation of symbol timing for carrier signals is very similar to the problem formulation for the baseband PAM signal.

6-3-2 Non-Decision-Directed Timing Estimation

A non-decision-directed timing estimate can be obtained by averaging the likelihood ratio $\Lambda(\tau)$ over the pdf of the information symbols, to obtain $\bar{\Lambda}(\tau)$, and then differentiating either $\bar{\Lambda}(\tau)$ or $\ln \bar{\Lambda}(\tau) = \bar{\Lambda}_L(\tau)$ to obtain the condition for the maximum-likelihood estimate $\hat{\tau}_{ML}$.

In the case of binary (baseband) PAM, where $I_n = \pm 1$ with equal probability, the average over the data yields

$$\bar{\Lambda}_L(\tau) = \sum_n \ln \cosh C y_n(\tau) \tag{6-3-7}$$

just as in the case of the phase estimator, Since $\ln \cosh x \approx \frac{1}{2}x^2$ for small x, the square-law approximation

$$\bar{\Lambda}_L(\tau) \approx \tfrac{1}{2}C^2 \sum_n y_n^2(\tau) \tag{6-3-8}$$

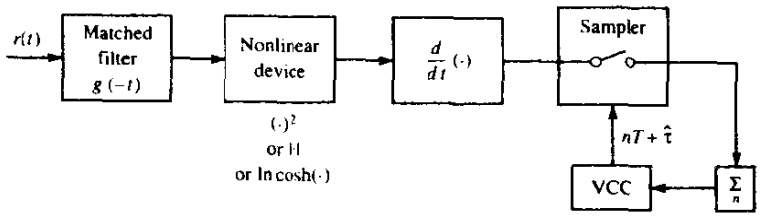
is appropriate for low signal-to-noise ratios. For multilevel PAM, we may approximate the statistical characteristics of the information symbols $\{I_n\}$ by the gaussian pdf, with zero mean and unit variance. When we average $\Lambda(\tau)$ over the gaussian pdf, the logarithm of $\bar{\Lambda}(\tau)$ is identical to $\bar{\Lambda}_L(\tau)$ given by (6-3-8). Consequently, the non-decision-directed estimate of τ may be obtained by differentiating (6-3-8). The result is an approximation to the ML estimate of the delay time. The derivative of (6-3-8) is

$$\frac{d}{d\tau}\sum_n y_n^2(\tau) = 2\sum_n y_n(\tau)\frac{dy_n(\tau)}{d\tau} = 0 \tag{6-3-9}$$

where $y_n(\tau)$ is given by (6-3-5).

An implementation of a tracking loop based on the derivative of $\bar{\Lambda}_L(\tau)$ given by (6-3-7) is shown in Fig. 6-3-2. Alternatively, an implementation of a

FIGURE 6-3-2 Non-decision-directed estimation of timing for binary baseband PAM.



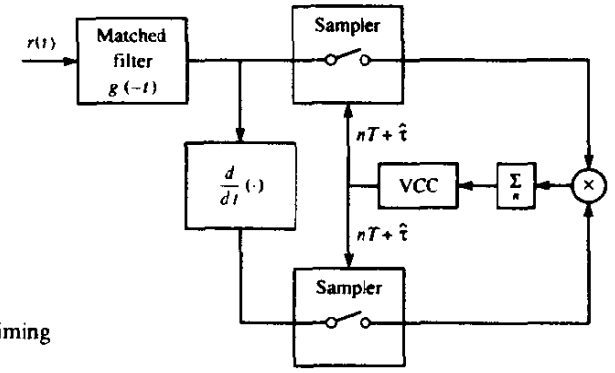


FIGURE 6-3-3 Non-decision-directed estimation of timing for baseband PAM.

tracking loop based on (6-3-9) is illustrated in Fig. 6-3-3. In both structures, we observe that the summation serves as the loop filter that drives the VCC. It is interesting to note the resemblance of the timing loop in Fig. 6-3-3 to the Costas loop for phase estimation.

Early–Late Gate Synchronizers Another non-decision-directed timing estimator exploits the symmetry properties of the signal at the output of the matched filter or correlator. To describe this method, let us consider the rectangular pulse $s(t)$, $0 \leq t \leq T$, shown in Fig. 6-3-4(a). The output of the filter matched to $s(t)$ attains its maximum value at time $t = T$, as shown in Fig. 6-3-4(b). Thus, the output of the matched filter is the time autocorrelation function of the pulse $s(t)$. Of course, this statement holds for any arbitrary pulse shape, so the approach that we describe applies in general to any signal pulse. Clearly, the proper time to sample the output of the matched filter for a maximum output is at $t = T$, i.e., at the peak of the correlation function.

In the presence of noise, the identification of the peak value of the signal is generally difficult. Instead of sampling the signal at the peak, suppose we sample early, at $t = T - \delta$ and late at $t = T + \delta$. The absolute values of the early samples $|y(m(T - \delta))|$ and the late samples $|y(m(T + \delta))|$ will be smaller (on the average in the presence of noise) than the samples of the peak value $|y(mT)|$. Since the autocorrelation function is even with respect to the optimum sampling time $t = T$, the absolute values of the correlation function at $t = T - \delta$ and $t = T + \delta$ are equal. Under this condition, the proper sampling

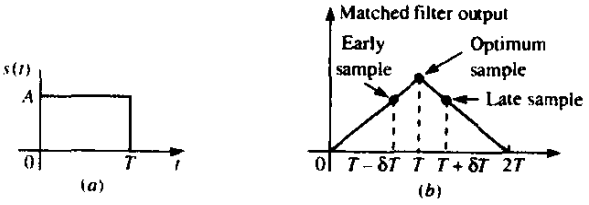


FIGURE 6-3-4 Rectangular signal pulse (a) and its matched filter output (b).

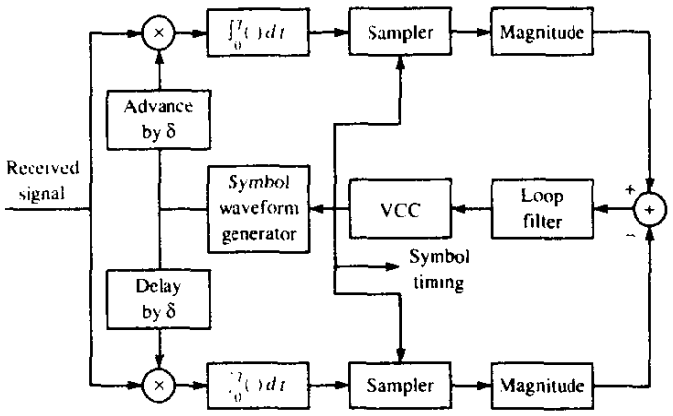


FIGURE 6-3-5 Block diagram of early–late gate synchronizer.

time is the midpoint between $t = T - \delta$ and $t = T + \delta$. This condition forms the basis for the *early–late gate symbol synchronizer.*

Figure 6-3-5 illustrates the block diagram of an early–late gate synchronizer. In this figure, correlators are used in place of the equivalent matched filters. The two correlators integrate over the symbol interval T, but one correlator starts integrating δ seconds early relative to the estimated optimum sampling time and the other integrator starts integrating δ seconds late relative to the estimated optimum sampling time. An error signal is formed by taking the difference between the absolute values of the two correlator outputs. To smooth the noise corrupting the signal samples, the error signal is passed through a lowpass filter. If the timing is off relative to the optimum sampling time, the average error signal at the output of the lowpass filter is nonzero, and the clock signal is either retarded or advanced, depending on the sign of the error. Thus, the smoothed error signal is used to drive a voltage-controlled clock (VCC), whose output is the desired clock signal that is used for sampling. The output of the VCC is also used as a clock signal for a symbol waveform generator that puts out the same basic pulse waveform as that of the transmitting filter. This pulse waveform is advanced and delayed and then fed to the two correlators, as shown in Fig. 6-3-5. Note that if the signal pulses are rectangular, there is no need for a signal pulse generator within the tracking loop.

We observe that the early–late gate synchronizer is basically a closed-loop control system whose bandwidth is relatively narrow compared to the symbol rate $1/T$. The bandwidth of the loop determines the quality of the timing estimate. A narrowband loop provides more averaging over the additive noise and, thus, improves the quality of the estimated sampling instants, provided that the channel propagation delay is constant and the clock oscillator at the transmitter is not drifting with time (or drifting very slowly with time). On the other hand, if the channel propagation delay is changing with time and/or the

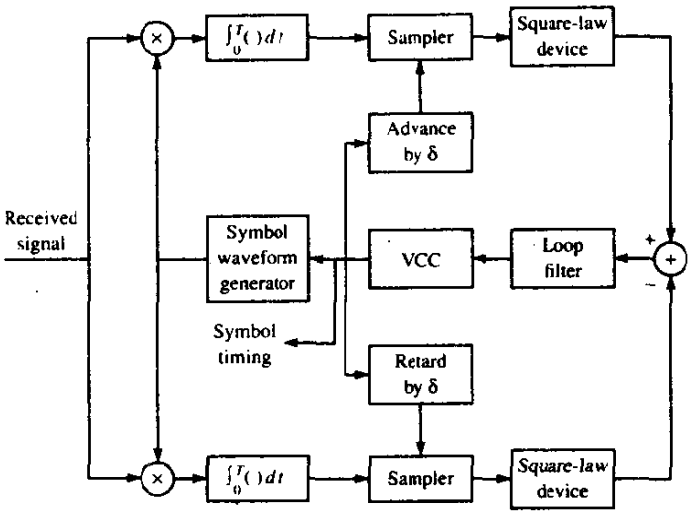


FIGURE 6-3-6 Block diagram of early–late gate synchronizer—an alternative form.

transmitter clock is also drifting with time then the bandwidth of the loop must be increased to provide for faster tracking of time variations in symbol timing.

In the tracking mode, the two correlators are affected by adjacent symbols. However, if the sequence of information symbols has zero mean, as is the case for PAM and some other signal modulations, the contribution to the output of the correlators from adjacent symbols averages out to zero in the lowpass filter.

An equivalent realization of the early–late gate synchronizer that is somewhat easier to implement is shown in Fig. 6-3-6. In this case the clock signal from the VCC is advanced and delayed by δ, and these clock signals are used to sample the outputs of the two correlators.

The early–late gate synchronizer described above is a non-decision-directed estimator of symbol timing that approximates the maximum-likelihood estimator. This assertion can be demonstrated by approximating the derivative of the log-likelihood function by the finite difference, i.e.,

$$\frac{d\bar{\Lambda}_L(\tau)}{d\tau} \approx \frac{\bar{\Lambda}_L(\tau+\delta) - \bar{\Lambda}_L(\tau-\delta)}{2\delta} \tag{6-3-10}$$

If we substitute for $\bar{\Lambda}_L(\tau)$ from (6-3-8) into (6-3-10), we obtain the approximation for the derivative as

$$\begin{aligned}\frac{d\bar{\Lambda}_L(\tau)}{d\tau} &= \frac{C^2}{4\delta}\sum_n [y_n^2(\tau+\delta) - y_n^2(\tau-\delta)] \\ &\approx \frac{C^2}{4\delta}\sum_n \left\{\left[\int_{T_0} r(t)g(t-nT-\tau-\delta)\,dt\right]^2 \right. \\ &\quad \left. - \left[\int_{T_0} r(t)g(t-nT-\tau+\delta)\,dt\right]^2\right\}\end{aligned} \tag{6-3-11}$$

But the mathematical expression in (6-3-11) basically describes the functions performed by the early–late gate symbol synchronizers illustrated in Figs 6-3-5 and 6-3-6.

6-4 JOINT ESTIMATION OF CARRIER PHASE AND SYMBOL TIMING

The estimation of the carrier phase and symbol timing may be accomplished separately as described above or jointly. Joint ML estimation of two or more signal parameters yields estimates that are as good and usually better than the estimates obtained from separate optimization of the likelihood function. In other words, the variances of the signal parameters obtained from joint optimization are less than or equal to the variance of parameter estimates obtained from separately optimizing the likelihood function.

Let us consider the joint estimation of the carrier phase and symbol timing. The log-likelihood function for these two parameters may be expressed in terms of the equivalent lowpass signals as

$$\Lambda_L(\phi, \tau) = \mathrm{Re}\left[\frac{1}{N_0}\int_{T_0} r(t)s_l^*(t; \phi, \tau)\,dt\right] \tag{6-4-1}$$

where $s_l(t; \phi, \tau)$ is the equivalent lowpass signal, which has the general form

$$s_l(t; \phi, \tau) = e^{-j\phi}\left[\sum_n I_n g(t - nT - \tau) + j\sum_n J_n w(t - nT - \tau)\right] \tag{6-4-2}$$

where $\{I_n\}$ and $\{J_n\}$ are the two information sequences.

We note that, for PAM, we may set $J_n = 0$ for all n, and the sequence $\{I_n\}$ is real. For QAM and PSK, we set $J_n = 0$ for all n and the sequence $\{I_n\}$ is complex-valued. For offset QPSK, both sequences $\{I_n\}$ and $\{J_n\}$ are nonzero and $w(t) = g(t - \frac{1}{2}T)$.

For decision-directed ML estimation of ϕ and τ, the log-likelihood function becomes

$$\Lambda_L(\phi, \tau) = \mathrm{Re}\left\{\frac{e^{j\phi}}{N_0}\sum_n [I_n^* y_n(\tau) + jJ_n^* x_n(\tau)]\right\} \tag{6-4-3}$$

where

$$\begin{aligned} y_n(\tau) &= \int_{T_0} r(t)g^*(t - nT - \tau)\,dt \\ x_n(\tau) &= \int_{T_0} r(t)w^*(t - nT - \tau)\,dt \end{aligned} \tag{6-4-4}$$

Necessary conditions for the estimates of ϕ and τ to be the ML estimates are

$$\frac{\partial\Lambda_L(\phi, \tau)}{\partial\phi} = 0, \qquad \frac{\partial\Lambda_L(\phi, \tau)}{\partial\tau} = 0 \tag{6-4-5}$$

It is convenient to define

$$A(\tau) + jB(\tau) = \frac{1}{N_0} \sum_n [I_n^* y_n(\tau) + jJ_n^* x_n(\tau)] \tag{6-4-6}$$

With this definition, (6-4-3) may be expressed in the simple form

$$\Lambda_L(\phi, \tau) = A(\tau) \cos \phi - B(\tau) \sin \phi \tag{6-4-7}$$

Now the conditions in (6-4-5) for the joint ML estimates become

$$\frac{\partial \Lambda(\phi, \tau)}{\partial \phi} = -A(\tau) \sin \phi - B(\tau) \cos \phi = 0 \tag{6-4-8}$$

$$\frac{\partial \Lambda(\phi, \tau)}{\partial \tau} = \frac{\partial A(\tau)}{\partial \tau} \cos \phi - \frac{\partial B(\tau)}{\partial \tau} \sin \phi = 0 \tag{6-4-9}$$

From (6-4-8), we obtain

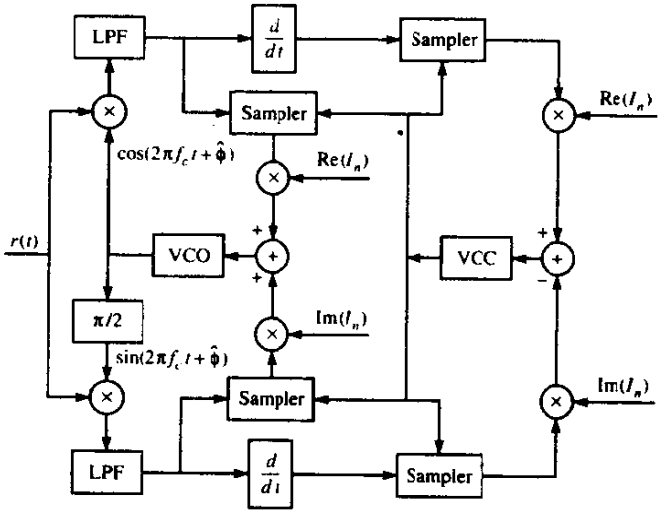
$$\hat{\phi}_{\mathrm{ML}} = -\tan^{-1} \left[\frac{B(\hat{\tau}_{\mathrm{ML}})}{A(\hat{\tau}_{\mathrm{ML}})} \right] \tag{6-4-10}$$

The solution to (6-4-9) that incorporates (6-4-10) is

$$\left[A(\tau) \frac{\partial A(\tau)}{\partial \tau} + B(\tau) \frac{\partial B(\tau)}{\partial \tau} \right]_{\tau = \hat{\tau}_{\mathrm{ML}}} = 0 \tag{6-4-11}$$

The decision-directed tracking loop for QAM (or PSK) obtained from these equations is illustrated in Fig. 6-4-1.

FIGURE 6-4-1 Decision-directed joint tracking loop for carrier phase and symbol timing in QAM and PSK.



Offset QPSK requires a slightly more complex structure for joint estimation of ϕ and τ. The structure is easily derived from (6-4-6)–(6-4-11).

In addition to the joint estimates given above, it is also possible to derive non-decision-directed estimates of the carrier phase and symbol timing, although we shall not pursue this approach.

We should also mention that one can combine the parameter estimation problem with the demodulation of the information sequence $\{I_n\}$. Thus, one can consider the joint maximum-likelihood estimation of $\{I_n\}$, the carrier phase ϕ, and the symbol timing parameter τ. Results on these joint estimation problems have appeared in the technical literature, e.g. Kobayashi (1971), Falconer (1976), and Falconer and Salz (1977).

6-5 PERFORMANCE CHARACTERISTICS OF ML ESTIMATORS

The quality of a signal parameter estimate is usually measured in terms of its bias and its variance. In order to define these terms, let us assume that we have a sequence of observations $[x_1\ x_2\ x_3\ \dots\ x_n] = \mathbf{x}$, with pdf $p(\mathbf{x} \mid \phi)$, from which we extract an estimate of a parameter ϕ. The bias of an estimate, say $\hat{\phi}(\mathbf{x})$, is defined as

$$\text{bias} = E[\hat{\phi}(\mathbf{x})] - \phi \tag{6-5-1}$$

where ϕ is the true value of the parameter. When $E[\hat{\phi}(\mathbf{x})] = \phi$, we say that the estimate is *unbiased*. The variance of the estimate $\hat{\phi}(x)$ is defined as

$$\sigma_{\hat{\phi}}^2 = E\{[\hat{\phi}(\mathbf{x})]^2\} - \{E[\hat{\phi}(\mathbf{x})]\}^2 \tag{6-5-2}$$

In general $\sigma_{\hat{\phi}}^2$ may be difficult to compute. However, a well-known result in parameter estimation (see Helstrom, 1968) is the Cramér–Rao lower bound on the mean square error defined as

$$E\{[\hat{\phi}(\mathbf{x}) - \phi]^2\} \geq \left[\frac{\partial}{\partial \phi} E[\hat{\phi}(\mathbf{x})]\right]^2 \Big/ E\left\{\left[\frac{\partial}{\partial \phi} \ln p(\mathbf{x} \mid \phi)\right]^2\right\} \tag{6-5-3}$$

Note that when the estimate is unbiased, the numerator of (6-5-3) is unity and the bound becomes a lower bound on the variance $\sigma_{\hat{\phi}}^2$ of the estimate $\hat{\phi}(\mathbf{x})$, i.e.,

$$\sigma_{\hat{\phi}}^2 \geq 1 \Big/ E\left\{\left[\frac{\partial}{\partial \phi} \ln p(\mathbf{x} \mid \phi)\right]^2\right\} \tag{6-5-4}$$

Since $\ln p(\mathbf{x} \mid \phi)$ differs from the log-likelihood function by a constant factor

independent of ϕ, it follows that

$$E\left\{\left[\frac{\partial}{\partial\phi}\ln p(\mathbf{x}\mid\phi)\right]^2\right\} = E\left\{\left[\frac{\partial}{\partial\phi}\ln\Lambda(\phi)\right]^2\right\}$$
$$= -E\left\{\frac{\partial^2}{\partial\phi^2}\ln\Lambda(\phi)\right\} \tag{6-5-5}$$

Therefore, the lower bound on the variance is

$$\sigma_{\hat{\phi}}^2 \geq 1\Big/E\left\{\left[\frac{\partial}{\partial\phi}\ln\Lambda(\phi)\right]^2\right\} = -1\Big/E\left[\frac{\partial^2}{\partial\phi^2}\ln\Lambda(\phi)\right] \tag{6-5-6}$$

This lower bound is a very useful result. It provides a benchmark for comparing the variance of any practical estimate to the lower bound. Any estimate that is unbiased and whose variance attains the lower bound is called an *efficient estimate*.

In general, efficient estimates are rare. When they exist, they are maximum-likelihood estimates. A well-known result from parameter estimation theory is that any ML parameter estimate is asymptotically (arbitrarily large number of observations) unbiased and efficient. To a large extent, these desirable properties constitute the importance of ML parameter estimates. It also known that an ML estimate is asymptotically gaussian-distributed [with mean ϕ and variance equal to the lower bound given by (6-5-6).]

In the case of the ML estimates described in this chapter for the two signal parameters, their variance is generally inversely proportional to the signal-to-noise ratio, or, equivalently, inversely proportional to the signal power multiplied by the observation interval T_0. Furthermore, the variance of the decision-directed estimates, at low error probabilities, are generally lower than the variance of non-decision-directed estimates. In fact, the performance of the ML decision-directed estimates for ϕ and τ attain the lower bound.

The following example is concerned with the evaluation of the Cramér–Rao lower bound for the ML estimate of the carrier phase.

Example 6-5-1

The ML estimate of the phase of an unmodulated carrier was shown in (6-2-11) to satisfy the condition

$$\int_{T_0} r(t)\sin(2\pi f_c t + \hat{\phi}_{\mathrm{ML}})\,dt = 0 \tag{6-5-7}$$

where

$$r(t) = s(t;\phi) + n(t)$$
$$= A\cos(2\pi f_c t + \phi) + n(t) \tag{6-5-8}$$

The condition in (6-5-7) was derived by maximizing the log likelihood function

$$\Lambda_L(\phi) = \frac{2}{N_0}\int_{T_0} r(t)s(t;\phi)\,dt \tag{6-5-9}$$

The variance of $\hat{\phi}_{\mathrm{ML}}$ is lower-bounded as

$$\begin{aligned}\phi^2_{\hat{\phi}_{\mathrm{ML}}} &\geq \left\{\frac{2A}{N_0}\int_{T_0} E[r(t)]\cos(2\pi f_c t + \phi)\,dt\right\}^{-1} \\ &\geq \left\{\frac{A^2}{N_0}\int_{T_0} dt\right\}^{-1} = \frac{N_0}{A^2 T_0} \\ &\geq \frac{N_0/2T_0}{\frac{1}{2}A^2} = \frac{N_0 B_{\mathrm{eq}}}{\frac{1}{2}A^2}\end{aligned} \tag{6-5-10}$$

The factor $1/2T_0$ is simply the (one-sided) equivalent noise bandwidth of the ideal integrator.

From this example, we observe that the variance of the ML phase estimate is lower-bounded as

$$\sigma^2_{\hat{\phi}_{\mathrm{ML}}} \geq 1/\gamma_L \tag{6-5-11}$$

where $\gamma_L = A^2/2N_0B_{\mathrm{eq}}$ is the loop SNR. This is also the variance obtained for the phase estimate from a PLL with decision-directed estimation. As we have already observed, non-decision-directed estimates do not perform as well due to losses in the nonlinearities required to remove the modulation, e.g., the squaring loss and the Mth-power loss.

Similar results can be obtained on the quality of the symbol timing estimates derived above. In addition to their dependence on the SNR, the quality of symbol timing estimates is a function of the signal pulse shape. For example, a pulse shape that is commonly used in practice is one that has a raised cosine spectrum (see Section 9-2). For such a pulse, the rms timing error (σ_τ) as a function of SNR is illustrated in Fig. 6-5-1, for both decision-directed and

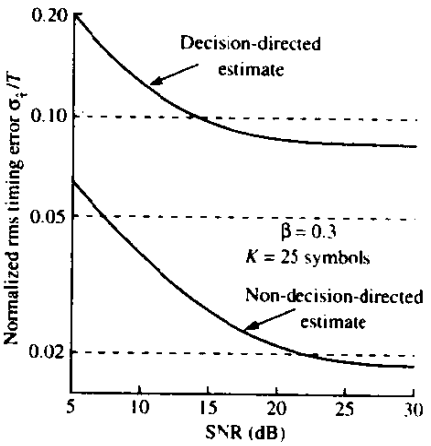


FIGURE 6-5-1 Performance of baseband symbol timing estimate for fixed signal and loop bandwidths. [*From* Synchronization Subsystems: Analysis and Design, *by L. Franks, 1983. Reprinted with permission of the author.*]

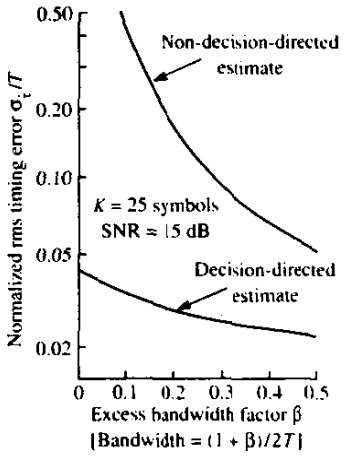


FIGURE 6-5-2 Performance of baseband symbol timing estimate for fixed SNR and fixed loop bandwidth. [*From* Synchronization Subsystems: Analysis and Design, *by L. Franks, 1983. Reprinted with permission of the author.*]

non-decision-directed estimates. Note the significant improvement in performance of the decision-directed estimate compared with the non-decision-directed estimate. Now, if the bandwidth of the pulse is varied, the pulse shape is changed and, hence, the rms value of the timing error also changes. For example, when the bandwidth of the pulse that has a raised cosine spectrum is varied, the rms timing error varies as shown in Fig. 6-5-2. Note that the error decreases as the bandwidth of the pulse increases.

In conclusion, we have presented the ML method for signal parameter estimation and have applied it to the estimation of the carrier phase and symbol timing. We have also described their performance characteristics.

6-6 BIBLIOGRAPHICAL NOTES AND REFERENCES

Carrier recovery and timing synchronization are two topics that have been thoroughly investigated over the past three decades. The Costas loop was invented in 1956 and the decision-directed phase estimation methods were described by Proakis *et al.* (1964) and by Natali and Walbesser (1969). The work on decision-directed estimation was motivated by earlier work of Price (1962a, b). Comprehensive treatments of phase-locked loops first appeared in the books by Viterbi (1966) and Gardner (1979). Books that cover carrier phase recovery and time synchronization techniques have been written by Stiffler (1971), Lindsey (1972), Lindsey and Simon (1973), and Meyr and Ascheid (1990).

A number of tutorial papers have appeared in IEEE journals on the PLL and on time synchronization. We cite, for example, the paper by Gupta (1975), which treats both analog and digital implementation of PLLs, and the paper by Lindsey and Chie (1981), which is devoted to the analysis of digital PLLs. In addition, the tutorial paper by Franks (1980) describes both carrier phase and symbol synchronization methods, including methods based on the maximum-likelihood estimation criterion. The paper by Franks is contained in a special

issue of the *IEEE Transactions on Communications* (August 1980) devoted to synchronization. The paper by Mueller and Muller (1976) describes digital signal processing algorithms for extracting symbol timing.

Application of the maximum-likelihood criterion to parameter estimation was first described in the context of radar parameter estimation (range and range rate). Subsequently, this optimal criterion was applied to carrier phase and symbol timing estimation as well as to joint parameter estimation with data symbols. Papers on these topics have been published by several researchers, including Falconer (1976), Mengali (1977), Falconer and Salz (1977), and Meyers and Franks (1980).

The Cramér–Rao lower bound on the variance of a parameter estimate is derived and evaluated in a number of standard texts on detection and estimation theory, such as Helstrom (1968) and Van Trees (1968). It is also described in several books on mathematical statistics, such as the book by Cramér (1946).

PROBLEMS

6-1 Prove the relation (6-1-7).

6-2 Sketch the equivalent realization of the binary PSK receiver in Fig. 6-1-1 that employs a matched filter instead of a correlator.

6-3 Suppose that the loop filter [see (6-2-14)] for a PLL has the transfer function

$$G(s) = \frac{1}{s + \sqrt{2}}$$

a Determine the closed-loop transfer function $H(s)$ and indicate if the loop is stable.

b Determine the damping factor and the natural frequency of the loop.

6-4 Consider the PLL for estimating the carrier phase of a signal in which the loop filter is specified as

$$G(s) = \frac{K}{1 + \tau_1 s}$$

a Determine the closed-loop transfer function $H(s)$ and its gain at $f = 0$.

b For what range of values of τ_1 and K is the loop stable?

6-5 The loop filter $G(s)$ in a PLL is implemented by the circuit shown in Fig. P6-5. Determine the system function $G(s)$ and express the time constants τ_1 and τ_2 in terms of the circuit parameters.

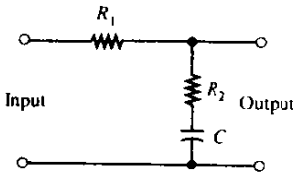


FIGURE P6-5

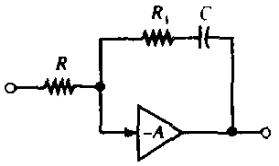


FIGURE P6-6

6-6 The loop filter $G(s)$ in a PLL is implemented with the active filter shown in Fig. P6-6. Determine the system function $G(s)$ and express the time constants τ_1 and τ_2 in terms of the circuit parameters.

6-7 Show that the early–late gate synchronizer illustrated in Fig. 6-3-5 is a close approximation to the timing recovery system illustrated in Fig. P6-7.

6-8 Based on a ML criterion, determine a carrier phase estimation method for binary on–off keying modulation.

6-9 In the transmission and reception of signals to and from moving vehicles, the transmitted signal frequency is shifted in direct proportion to the speed of the vehicle. The so-called *Doppler frequency shift* imparted to a signal that is received in a vehicle traveling at a velocity v relative to a (fixed) transmitter is given by the formula

$$f_D = \pm\frac{v}{\lambda}$$

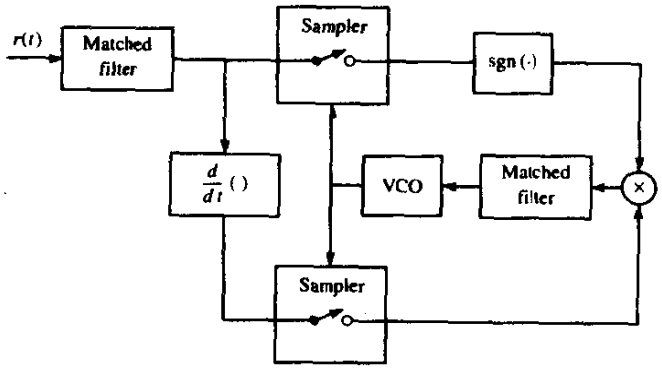
where λ is the wavelength, and the sign depends on the direction (moving toward or moving away) that the vehicle is traveling relative to the transmitter. Suppose that a vehicle is traveling at a speed of 100 km/h relative to a base station in a mobile cellular communication system. The signal is a narrowband signal transmitted at a carrier frequency of 1 GHz.

a Determine the Doppler frequency shift.

b What should be the bandwidth of a Doppler frequency tracking loop if the loop is designed to track Doppler frequency shifts for vehicles traveling at speeds up to 100 km/h?

c Suppose the transmitted signal bandwidth is 2 MHz centered at 1 GHz.

FIGURE P6-7



Determine the Doppler frequency spread between the upper and lower frequencies in the signal.

6-10 Show that the mean value of the ML estimate in (6-2-38) is ϕ, i.e., that the estimate is unbiased.

6-11 Determine the pdf of the ML phase estimate in (6-2-38).

6-12 Determine the ML phase estimate for offset QPSK.

6-13 A single-sideband PAM signal may be represented as

$$u_m(t) = A_m[g_I(t)\cos 2\pi f_c t - \hat{g}_I(t)\sin 2\pi f_c t]$$

where $\hat{g}_I(t)$ is the Hilbert transform of $g_I(t)$ and A_m is the amplitude level that conveys the information. Demonstrate mathematically that a Costas loop can be used to demodulate the SSB PAM signal.

6-14 A carrier component is transmitted on the quadrature carrier in a communication system that transmits information via binary PSK. Hence, the received signal has the form

$$r(t) = \pm\sqrt{2P_s}\cos(2\pi f_c t + \phi) + \sqrt{2P_c}\sin(2\pi f_c t + \phi) + n(t)$$

where ϕ is the carrier phase and $n(t)$ is AWGN. The unmodulated carrier component is used as a pilot signal at the receiver to estimate the carrier phase.

a Sketch a block diagram of the receiver, including the carrier phase estimator.

b Illustrate mathematically the operations involved in the estimation of the carrier phase ϕ.

c Express the probability of error for the detection of the binary PSK signal as a function of the total transmitted power $P_T = P_s + P_c$. What is the loss in performance due to the allocation of a portion of the transmitted power to the pilot signal? Evaluate the loss for $P_c/P_T = 0.1$.

6-15 Determine the signal and noise components at the input to a fourth-power ($M = 4$) PLL that is used to generate the carrier phase for demodulation of QPSK. By ignoring all noise components except those that are linear in the noise $n(t)$, determine the variance of the phase estimate at the output of the PLL.

6-16 The probability of error for binary PSK demodulation and detection when there is a carrier phase error ϕ_e is

$$P_2(\phi_e) = Q\left(\sqrt{\frac{2\mathcal{E}_b}{N_0}\cos^2\phi_e}\right)$$

Suppose that the phase error from the PLL is modeled as a zero-mean gaussian random variable with variance $\sigma_\phi^2 \ll \pi$. Determine the expression for the average probability of error (in integral form).