A BPSK/QPSK Timing-Error Detector for Sampled Receivers

FLOYD M. GARDNER, FELLOW, IEEE

Abstract—**A simple algorithm for detection of timing error of a synchronous, band-limited, BPSK or QPSK data stream is proposed. The algorithm requires only two samples per symbol for its operation. One of the two samples is also used for the symbol decision. Derivation of the *s*-curve reveals a sinusoidal shape.**

I. Introduction

SAMPLED implementations of receivers for digital data signals are growing more popular as components—notably digital signal processors—improve in capability. There is a need for sampled algorithms to replace the continuous-time methods that have predominated heretofore.

This paper introduces an algorithm for timing error detection in a receiving modem. Algorithm operations are simple and only two samples of the signal are required for each data symbol. Moreover, one of the two samples also serves for the symbol strobe (i.e., the sample on which the symbol decision is made).

The algorithm is intended for synchronous, binary, baseband signals and for BPSK or QPSK (balanced, nonstaggered) passband signals, with approximately 40–100 percent excess bandwidth.

Other sampled timing algorithms have been presented earlier. Mueller and Müller [1] wrote a classic explanation of timing recovery based on just one sample per symbol and requiring decision-directed operations. Note that correct decisions in a carrier system depend upon prior acquisition of carrier phase.

By contrast, the algorithm introduced in this paper is not decision directed. Furthermore, clock recovery is quite independent of carrier phase.

Other papers [2]–[5] have followed the lead of [1]. A different scheme, dubbed the *wave difference method* (WDM), was proposed in [6]. In essence, the WDM finds the average location of zero-slope of the received, filtered signal pulses. Numerous samples per symbol appear to be required for the original method of [6].

Reference [7] extends the WDM, shows how to reduce it to two samples per symbol, suggests implementation details, and provides analyses of performance. Neither of the two sample points in [7] coincides with the decision strobe point. Either four samples per symbol must be taken or else (as in [7]) the two samples per symbol must be interpolated to an effective four samples per symbol. (Interpolation in [7] is accomplished with a digital delay of a quarter-symbol interval.) Both [6] and [7] treat clock recovery only for baseband signals.

The method introduced in this paper, although developed independently, has resemblances to the WDM of [7]. Its major points of difference are: only two samples per symbol are employed, without explicit interpolation; one sample coincides with the decision instant; and carrier signals are handled as well as baseband signals.

References [1] and [7] contain informative summaries of analog clock-recovery methods, with references.

This paper is organized as follows. A model of a receiving modem is shown in Section II and some characteristics of the new algorithm are stated in Section III. A baseband version of the algorithm is developed in Section IV. Derivation of the *s*-curve (sinusoidal for a band-limited signal) is relegated to Appendix A, while Appendix B demonstrates that the algorithm is independent of carrier phase.

II. Receiver Model

Refer to Fig. 1 for a block diagram of a typical *I*–*Q* receiving modem. A passband signal is demodulated to baseband in a pair of quadrature-driven mixers. Phase of the local carrier must be adjusted to agree with that of the signal. The necessary carrier-recovery branch is omitted from the diagram and is irrelevant to the clock algorithm and discussion.

Data filters follow the mixers; they perform receiver filtering to shape signal pulses, minimize noise, and suppress unwanted mixer products.

Our interest is in sampled receivers. We do not specify the sampling point other than to say that the filter outputs are available only in sampled form as the pair of real sequences $\{y_I(\ )\}$ and $\{y_Q(\ )\}$. Timing information must be retrieved from these sequences.

Symbols are transmitted synchronously, spaced by the time interval T. Each sequence will have two samples per symbol interval and the samples will be time-coincident between the sequences. One sample occurs at the data strobe time and the other sample occurs midway between data strobe times.

The index r is used to designate symbol number. It is convenient to denote the strobe values of the rth symbol as $y_I(r)$ and $y_Q(r)$. As a formalism, we denote the values of the pair of samples lying midway between the $(r-1)$th and the rth strobes as $y_I(r-1/2)$ and $y_Q(r-1/2)$.

A timing error detector operates upon samples and generates one error sample $u_t(r)$ for each symbol. The actions of that detector are the main focus of this paper.

The error sequence is smoothed by a loop filter and then used to adjust a timing error corrector. This paper treats only the detector and is not concerned with the loop filter nor with the error corrector.

III. Detector Characteristics

A detector algorithm of the form

$$u_t(r) = y_I(r-1/2)[y_I(r) - y_I(r-1)] + y_Q(r-1/2)[y_Q(r) - y_Q(r-1)] \quad (1)$$

is derived in the sequel. This algorithm is suitable for both tracking and acquisition modes of operation. It is proven in Appendix B that $u_t(r)$ is independent of carrier phase, so that

Paper approved by the Editor for Synchronization and Spread Spectrum of the IEEE Communications Society. Manuscript received September 25, 1985; revised December 23, 1985.

The author is at 1755 University Avenue, Palo Alto, CA 94301.

IEEE Log Number 8607993.

0090-6778/86/0500-0423$01.00

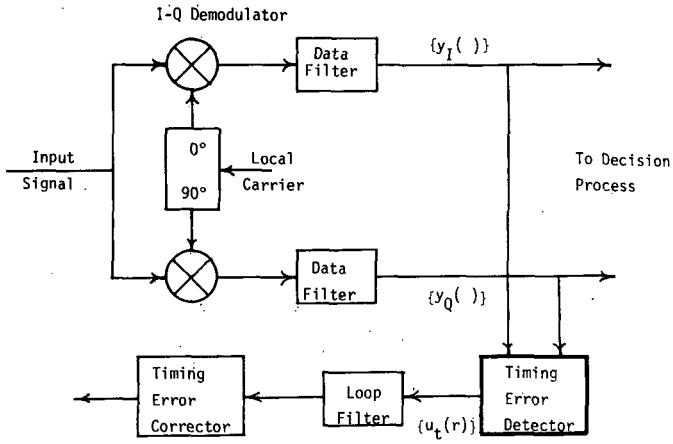


Fig. 1. Typical modem (simplified block diagram).

timing lock can be achieved without depending upon prior carrier phase lock.

If BPSK is employed, and if the modulation is in the I channel, then the Q channel terms contribute no information once carrier phase lock has been achieved. Only noise would come from the Q channel after carrier lock. Therefore, the full algorithm of (1) would be needed for best BPSK timing acquisition, but the algorithm should be reduced to just the I-channel terms after carrier lock is achieved. Similarly, a baseband signal consists of just one channel, carrier acquisition is not an issue, and so only the $\{y_I(\ )\}$ terms are applicable.

If a QPSK signal is being processed, then both I and Q arms contribute useful information after carrier lock; both parts of (1) should be retained for QPSK clock tracking.

An analysis of the detector characteristic of (1) is given in Appendix A, where it is shown that a band-limited signal produces a sinusoidal s-curve. Once the data-pulse shape is known, that analysis can be applied to determine the gain of the detector.

The method proposed here is thought to be applicable for excess bandwidths of the signal of approximately 40–100 percent; this range is representative of satellite communications. It is known that the proposed algorithm has rapidly deteriorating noise performance with narrower bandwidth, partly because the efficiency of clock regeneration falls off (as will be further noted in Appendix A) and partly because of increasing self noise. This feature is inherent to regenerators with quadratic nonlinearities. A different algorithm, based on a harder nonlinearity, would be needed for narrow bandwidths. (But we think the independence of carrier phase is unique to a square-law nonlinearity.)

A physical explanation can be ascribed to (1). The detector samples the data stream midway between strobe locations in each of the I and Q channels. If there is a transition between symbols, the average midway value should be zero, in the absence of timing error. A timing error gives a nonzero sample whose magnitude depends upon the amount of error, but either slope is equally likely at the midway point so there is no direction information in the sample alone.

To sort out these different possibilities, the algorithm examines the two strobe values to either side of the midway sample. If there is no transition, the strobe values are the same, their difference is zero, and so the midway sample is rejected. (No timing information is available in the absence of a transition.)

If a transition is present, the strobe values will be different; the difference between them will provide slope information. The product of the slope information and the midway sample provides timing-error information.

It may be worthwhile to use the signs of the strobe values instead of the actual values. That eliminates the effects of much noise. If all data filtering has been performed prior to the strobe point, then the sign of the strobe value is the optimum hard decision on the symbol and the algorithm effectively becomes decision directed. This expedient is known to improve tracking capability. (But acquisition performance may suffer in a decision-directed operation.) Note that use of the strobe signs, instead of actual values, eliminates the need for actual multiplications in the algorithm—an attractive feature for digital processors.

The decision-directed version of the algorithm will be recognized as very similar to the digital transition tracking loop of Lindsey and Simon [8].

We can see a source of self noise in either version of this algorithm. If the excess bandwidth is less than 100 percent, the zero crossings of data transitions do not all lie midway between the strobe points. There is a scatter of crossing points, centered on the midway point. The average location is correct, but any individual trajectory can depart from the average, causing self noise.

One last comment is appropriate in this section: The algorithm takes its information from three different sample points in order to produce one timing-error point. There is a delay in computing the error sample. The timing can be adjusted within the three-point span of the algorithm, so the timing error at the last point in the span does not necessarily remain the same as the error at the first point. Equivalently, the algorithm contains memory. These features must be taken into account in the analysis of the tracking loop.

IV. Derivation of Algorithm

This section develops the timing detector by physical reasoning from data waveforms. Past experience with timing recovery in analog systems is used as a guide for the all-digital detector.

Waveform Approach

We start out with the problem of recovering timing information from a baseband data stream. To that end, we postulate the existence of an equivalent analog, baseband signal underlying the digital sequence that actually occurs in the processor. It is easier to visualize operations on the continuous signal than on the discrete sequence. Any zero-memory operations performed on the continuous signal will commute with sampling, so that equivalent results are obtained by performing the same operations on the discrete sequence.

A set of typical waveforms is shown in Fig. 2. Line A shows symbol boundaries (of width T seconds) and line B shows locations of the strobes: one per symbol in the center of each symbol interval.

Line C shows a hypothetical baseband signal $x(t)$. Examination will show that the Nyquist-1 criterion is met and that the signal has been moderately band limited. Actually, the signal pulses used for the illustration were raised cosines, which are very close to band-limited Nyquist pulses with 100 percent excess bandwidth.

Simple Rectifier: One well-proven method of regenerating a clock wave from data stream is to pass the stream through a rectifier [9]. We use that expedient as a starting point.

A square-law rectifier has several advantages:

- Its noise performance is near-optimum, especially at low SNR [9].
- A pure sinusoidal input results in a pure sinusoidal output at double the frequency.
- A square law is almost the only nonlinearity that is mathematically tractable.
- Square-law rectifiers in both I and Q arms will recover a clock wave independent of carrier phase—an important consideration for acquisition.

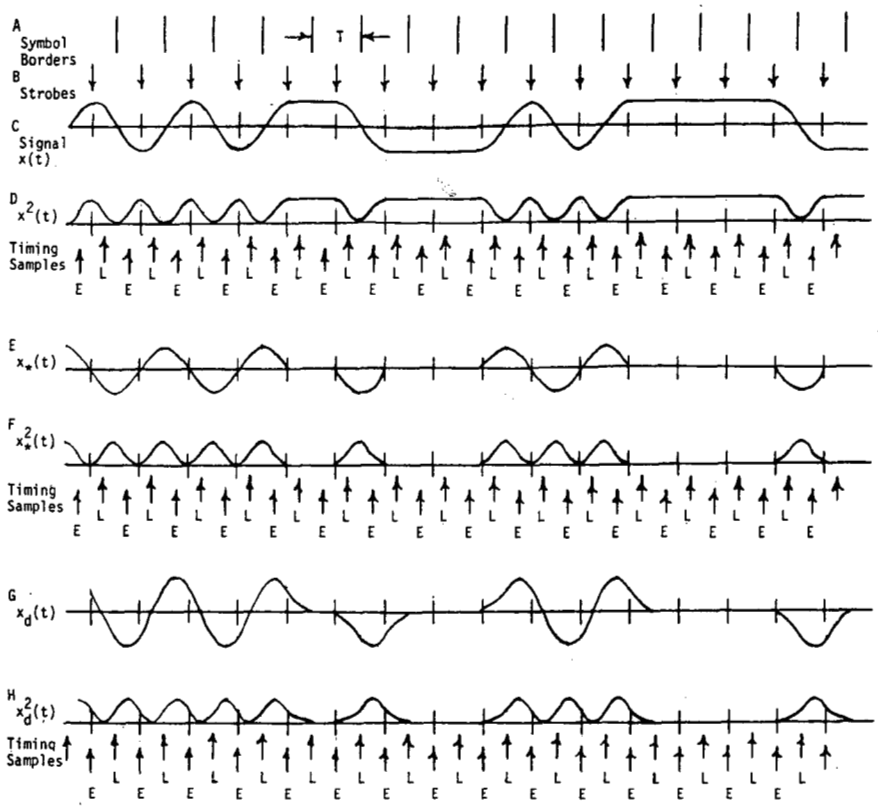


Fig. 2. Timing waveforms.

We will start the derivation with a square-law rectifier.

Line D of Fig. 2 shows the data stream after square-law rectification. Evident in the picture are a dc component (an inevitable output from a rectifier) plus a double-frequency component at the symbol rate. This *protoclock* component is equivalent to a pure sinewave at the symbol frequency, interrupted by gaps when there are no transitions in the data. An analog clock-recovery scheme would employ a narrow-band filter or a phase-lock loop to extract the desired clock line and reject the various disturbances.

The same approach is feasible in a digital system, but is not very convenient. Proper operation with a filter or PLL scheme would necessitate good reconstruction of the protoclock and that cannot be done with just two samples per symbol. (Two samples per cycle of a sinewave is the upper limit of the sampling theorem; the protoclock has a bandwidth wider than $1/T$ and cannot be reconstructed from samples at a rate of $2/T$.)

Our objective is not to reconstruct the protoclock, but to determine timing error—which does not inherently require protoclock reconstruction.

Examining line D, it can be seen that samples taken at locations of $\pm 1/4T$ away from the strobe point fall at equal-amplitude, but opposite-slope points on the useful portions of the protoclock. When timing is correct, the difference between those two samples—labeled E for early and L for late—will be zero (in those locations where a data transition exists). If the signal is delayed with respect to the sampling, then the value of one sample increases and the other decreases. The difference between sample values is a measure of timing error.

To formalize the method, assign an index r to each timing sample, so that the two timing samples in the rth symbol interval are denoted $E(r)$ and $L(r)$. This introductory algorithm would be

$$u_t(r) = E(r) - L(r-1)$$
$$= x^2(\tau + (r - 1/4)T) - x^2(\tau + (r - 5/4)T) \quad (2)$$

whereas the rth strobe is taken at $t = rT + \tau$, and τ is the timing shift from desired delay.

Note that in the absence of additive noise, and for the idealized waveforms used in the example, that $u_t(\ )$ will be zero when transitions are missing. That feature avoids significant self noise.

A slightly different algorithm, $E(r) - L(r)$, could have been employed instead. It would also provide a measure of the timing error and a successful loop could be based upon it. However, a large error sample is generated by this algorithm for many locations where there are no transitions. It is true that an opposite-sign error-sample is certain to be generated later and that the two large samples cancel out in the long-term tracking. However, in the short term, the large error samples contribute to self noise. For that reason, the selected algorithm is superior to the alternative. The same choice, with the same self-noise consequences, will reappear in subsequent developments. Without further comment, we will select the sample pair giving the best rejection of self noise.

The algorithm of (2) is workable. From experience with continuous systems [10] we can anticipate that satisfactory performance could be attained in typical communications links. However, it has two deficiencies that should be remedied.

- Although it needs only two samples per symbol, neither of those samples is the data-strobe sample. Therefore, the postfilter operations require at least three samples per symbol, thereby placing a greater computing burden on the data filters.
- No self noise arises with the idealized waveforms of the illustration. But that happy condition occurs only because the illustrative waveforms are time-limited. When band-limited, time-extended waveforms are used, then self noise will arise to some extent.

Let us attack the self noise first.

It is well established [11] that a suitable prefilter in front of the rectifier, plus suitable postfiltering, can suppress self noise completely. That is the next approach examined.

Prefilter Method: Line E shows the output $x_*(t)$ of a prefilter which has $x(t)$, from line C, as its input. For time-limited pulses [12], $x_*(t)$ is simply the derivative of $x(t)$. For band-limited pulses, the spectrum of $x_*(t)$ is band limited and symmetric about the Nyquist frequency $1/2T$ (See [11]).

A protoclock is generated by rectifying $x_*(t)$, as illustrated in line F. For the particular, idealized waveforms used in Fig. 2, it turns out that Line F is simply the inversion and level shift of line D. The same arguments on sampling and the same timing algorithm are appropriate.

In the absence of a postfilter, it is not apparent that the prefilter affords self-noise relief for band-limited waveforms. A more detailed analysis would be needed to demonstrate that feature. Instead, we merely note that the idealized prefilter for the idealized illustrative waveform has not relieved the shift of the timing samples from the strobe location.

A differentiator, for time-limited pulses, or an optimum prefilter, for band-limited pulses, it is likely to be rather complicated and would add unduly to the computation burden. It has been shown [13] that fairly crude approximations to ideal prefilters provide substantial improvement of self noise. Therefore, we compromise optimality in the interest of simplicity and restrict the method to very simple prefilters.

Delay Differencing: Let the prefilter or differentiator be approximated by the operation

$$x_d(t) = x(t) - x(t - t_d) \tag{3}$$

where t_d is a suitably chosen delay time. In particular, if $t_d = T/2$, then the average delay of x_d is $T/4$—which is the time shift between the strobe point and the timing samples that has arisen in the two previous approaches.

Line G shows the result if the signal of line C is delay-differenced with a delay of $t_d = T/2$. Comparison with line E shows similar, but not identical waveforms.

The delay-differenced waveform is passed through a square-law rectifier, yielding line H, for the example waveform. Once again, the protoclock appears and timing-error information can be retrieved by the algorithm

$$u_t(r) = E(r) - L(r-1). \tag{4}$$

However, now observe that $E(r)$ coincides with the strobe time for the rth symbol; the delay in the differencing network shifted the time for the timing samples and eliminated the need for a third sample. One timing sample now coincides with the data strobe and the other falls midway between data strobes. (The strobe locations are carried on each waveform abscissa as vertical tick marks in Fig. 2.)

In this roundabout manner we have devised a timing-error detector algorithm that requires only two samples per symbol (one sample coincident with the data strobe), that has some self-noise rejection properties, and that can be performed with minimal computing burden (one subtraction and one squaring per symbol interval). Now let us see what improvements are possible.

Formal Reductions

In this section we perform formal algebraic manipulations on the algorithm expression and discover that there are elements hidden in the algorithm that contribute no useful output. Since their presence could only generate self noise, they are eliminated and a stripped algorithm is derived instead.

Start with

$$x_d(t) = x(t) - x(t - T/2) \tag{5}$$

$$x_d^2(t) = x^2(t) + x^2(t - T/2) - 2x(t)x(t - T/2). \tag{6}$$

Upon sampling at $t = rT + \tau$ and $rT + \tau - T/2$, we obtain

$$\begin{aligned} E(r) &= x_d^2(rT+\tau) \\ &= x^2(\tau + rT) + x^2(\tau + (r - 1/2)T) \\ &\quad - 2x(\tau + rT)x(\tau + (r - 1/2)T) \\ L(r-1) &= x^2(\tau + (r - 1/2)T) + x^2(\tau + (r-1)T) \\ &\quad - 2x(\tau + (r - 1/2)T)x(\tau + (r-1)T). \end{aligned} \tag{7}$$

Let the algorithm be

$$u_t(r) = L(r-1) - E(r). \tag{8}$$

The reversal of sign has no significance in the formal manipulations or in the processor's computation burden, but assures negative slope at the tracking point of the detector output. See Appendix A.

Both elements of (8) contain a term $x^2(\tau + (r - 1/2)T)$; the subtraction cancels those terms. Collecting terms after the cancellation, the algorithm is composed of the elements

$$\begin{aligned} u_t(r) &= x^2(\tau + (r-1)T) - x^2(\tau + rT) \\ &\quad + 2x(\tau + (r - 1/2)T)\{x(\tau + rT) - x(\tau + (r-1)T)\}. \end{aligned} \tag{9}$$

Useful output of the algorithm is the average over many samples, not the value of an isolated sample. Denote $U_t(\tau) = \text{Avg}_r\ u_t(r)$ as the average over many samples, so that

$$\begin{aligned} U_t(\tau) &= \text{Avg}\ \{x^2(\tau + (r-1)T\} - \text{Avg}\ \{x^2(\tau + rT)\} \\ &\quad + 2\ \text{Avg}\ \{x(\tau + (r - 1/2)T)(x(\tau + rT) - x(\tau + (r-1)T))\}. \end{aligned} \tag{10}$$

The first two terms must be equal; the ensemble average cannot depend upon the index r because the underlying signal is cyclostationary. Therefore, those averages must cancel. The presence of the $x^2(\ )$ terms does not contribute to the useful average output; we must suspect them as potential sources of self noise. Therefore, let us eliminate them.

The remaining terms are of the form

$$\begin{aligned} u_t(r) &= x(\tau + (r - 1/2)T)\{x(\tau + rT) - x(\tau + (r-1)T)\} \\ &= x(r - 1/2)\{x(r) - x(r-1)\}. \end{aligned} \tag{11}$$

This is the proposed timing-detector algorithm for real, baseband signals. Two such computations, one each from the I and from the Q channel, are added when processing a demodulated carrier signal. The sum is the algorithm of (1).

APPENDIX A

DERIVATION OF *S*-CURVE

Satisfactory design of the timing loop requires that the detector characteristic—average output versus timing error τ—be known. That is derived for the baseband algorithm of (11) in this Appendix. A phase-independent extension to the I–Q algorithm of (1) is demonstrated in Appendix B.

Let the underlying, time-continuous signal be a PAM stream with the format

$$x(t) = \Sigma a_p g(t - pT) \tag{12}$$

where this summation, and those to follow, is assumed to be doubly infinite.

The sequence $\{a_p\}$ is taken from a binary library where $a_p = \pm 1$. We will assume that the a_p have zero mean and are uncorrelated. That is,

$$E(a_p) = 0$$

$$E(a_p a_q) = \delta_{pq} E(a_p^2) = \delta_{pq}. \tag{13}$$

The function $g(t)$ is the shape of the filtered signal pulse. It has Fourier transform $G(f)$.

Apply (12) to (11) to obtain

$$u_t(r) = \sum_p a_p g(\tau + (r - 1/2)T - pT) \times \left(\sum_q a_q g(\tau + rT - qT) - \sum_s a_s g(\tau + (r-1)T - sT) \right). \tag{14}$$

Perform the indicated multiplications and collect terms, thereby obtaining double sums containing products $a_p a_q$ and $a_p a_s$. Take the expectation over the data ensemble to arrive at the average output of the detector:

$$\begin{aligned} U_t(\tau) &= E\{u_t(r)\} \\ &= \sum_p g(\tau + (r-1/2)T - pT)[g(\tau + rT - pT) \\ &\quad - g(\tau + (r-1)T - pT)] \\ &= \sum_q g(\tau + (q-1/2)T) g(\tau + qT) \\ &\quad - \sum_q g(\tau + (q-1/2)T) g(\tau + (q-1)T) \end{aligned} \tag{15}$$

where $q = r - p$.

Consider a partial term of the form

$$\alpha(\tau) = g(\tau + (q-1/2)T) g(\tau + qT) = \int_{-\infty}^{\infty} A(f) e^{j2\pi f}\, df. \tag{16}$$

Then

$$\begin{aligned} A(f) &= \int g(\tau + (q-1/2)T) g(\tau + qT) e^{-j2\pi f\tau}\, d\tau \\ &= \int g(\tau + (q-1/2)T) e^{-j2\pi f\tau} \int G(\nu) e^{j2\pi\nu(\tau + qT)}\, d\nu\, d\tau \\ &= \int G(\nu) e^{j2\pi\nu qT} \int g(\tau + (q-1/2)T) e^{-j2\pi\tau(f-\nu)}\, d\tau\, d\nu \end{aligned} \tag{17}$$

where all integrals are over double-infinite limits until otherwise stated.

Let $\tau' = \tau + (q - 1/2)T$, whereby

$$\begin{aligned} A(f) &= \int G(\nu) e^{j2\pi\nu qT} \int g(\tau') e^{-j2\pi(\tau' - (q-(1/2))T)(f-\nu)}\, d\tau'\, d\nu \\ &= e^{j2\pi f(q-(1/2))T} \int G(\nu) G(f-\nu) e^{j\pi\nu T}\, d\nu. \end{aligned} \tag{18}$$

Return to the time domain via

$$\begin{aligned} \alpha(\tau) &= \int A(f) e^{j2\pi f\tau}\, df \\ &= \int G(\nu) e^{j\pi\nu T} \int G(f-\nu) e^{j2\pi f\tau} e^{j2\pi f(q-(1/2))T}\, df\, d\nu. \end{aligned} \tag{19}$$

We really want the infinite sum

$$\sum_q \alpha(\tau, q) = \int G(\nu) e^{j\pi\nu T} \cdot \int G(f-\nu) e^{j2\pi f\tau} e^{-j\pi fT} \sum_q e^{j2\pi fqT}\, df\, d\nu. \tag{20}$$

It can be demonstrated that

$$T \sum_{q=-\infty}^{\infty} e^{j2\pi fqT} = \sum_{m=-\infty}^{\infty} \delta(f - m/T) \tag{21}$$

so

$$\begin{aligned} \sum_q \alpha(\tau, q) &= \frac{1}{T} \int G(\nu) e^{j\pi\nu T} \sum_m G\left(\frac{m}{T} - \nu\right) e^{j2\pi m\tau/T} e^{-j\pi m}\, d\nu \\ &= \frac{1}{T} \sum_{m=-1}^{1} (-1)^m e^{j2\pi m\tau/T} \cdot \int_{-1/T}^{1/T} G(\nu) G\left(\frac{m}{T} - \nu\right) e^{j\pi\nu T}\, d\nu. \end{aligned} \tag{22}$$

In the last line we have imposed the restriction $G(f) = 0$ for $|f| \geq 1/T$ to take account of band limiting and to simplify the analysis. That restriction assures that there is no nonzero overlap between $G(f)$ and $G((m/T) - f)$ for $|m| > 1$. The only contributing values of m are 0 and ± 1.

In similar manner, define

$$\beta(\tau) = g(\tau + (q-1/2)T) g(\tau + (q-1)T)$$

and, undertaking the same manipulations already shown for $\alpha(\tau)$, we obtain

$$\sum_q \beta(\tau, q) = \frac{1}{T} \sum_m (-1)^m e^{j2\pi m\tau/T} \cdot \int G(\nu) G\left(\frac{m}{T} - \nu\right) e^{-j\pi\nu T}\, d\nu. \tag{23}$$

The quantity of interest is

$$\begin{aligned} U_t(\tau) &= \sum_q \{\alpha(\tau) - \beta(\tau)\} \\ &= 2j(1/T) \sum_m (-1)^m e^{j2\pi m\tau/T} \cdot \int G(f) G\left(\frac{m}{T} - f\right) \sin \pi fT\, df. \end{aligned} \tag{24}$$

Let us examine this expression separately for each of the three different contributory values of m.

$m = 0$:

$$(-1)^0 = 1; \quad e^{j0} = 1.$$

If $g(t)$ is real, then

$$\int G(f) G(-f) \sin \pi fT\, df = 0$$

because $G(f)G(-f)$ is even and $\sin \pi fT$ is odd; the integral of their product, over symmetric limits, must vanish.

$m = \pm 1$:

$$(-1)^{\pm 1} = -1; \quad e^{j2\pi m\tau/T} = e^{\pm j2\pi\tau/T}$$

so the surviving terms are

$$\begin{aligned} U_t(\tau) = -2j(1/T) \Bigg(& e^{j2\pi\tau/T} \int_0^{1/T} G(f) \cdot G\left(\frac{1}{T} - f\right) \sin \pi fT\, df \\ & + e^{-j2\pi\tau/T} \int_{-1/T}^{0} G(f) \cdot G\left(-\frac{1}{T} - f\right) \sin \pi fT\, df \Bigg) \end{aligned} \tag{25}$$

where the asymmetric limits come about because one of the shifted $G(\ )$ terms vanishes in a region where $G(f)$ itself is nonzero.

This expression is valid for any band-limited $g(t)$; it can be evaluated by substituting $G(f)$ and working out the integrals of (25).

Better insight is gained by resolving $G(f)$ into its even and odd components. For $g(t)$ a real function we find

$$G(f) = G_e(f) + jG_0(f) \tag{26}$$

where G_e and G_0 are both purely real, with properties

$$G_e(f) = G_e(-f)$$

$$G_0(f) = -G_0(-f).$$

Further, define the shorthand notation

$$G_{R+} = \int_0^{1/T} \left(G_e(f) G_e\left(\frac{1}{T} - f\right) - G_0(f) G_0\left(\frac{1}{T} - f\right)\right) \sin\ \pi f T\ df$$

$$G_{I+} = \int_0^{1/T} \left(G_0(f) G_e\left(\frac{1}{T} - f\right) + G_e(f) G_0\left(\frac{1}{T} - f\right)\right) \sin\ \pi f T\ df$$

$$G_{R-} = \int_{-1/T}^{0} \left(G_e(f) G_e\left(-\frac{1}{T} - f\right) - G_0(f) G_0\left(-\frac{1}{T} - f\right)\right) \sin\ \pi f T\ df$$

$$G_{I-} = \int_{-1/T}^{0} \left(G_0(f) G_e\left(-\frac{1}{T} - f\right) + G_e(f) G_0\left(-\frac{1}{T} - f\right)\right) \sin\ \pi f T\ df. \tag{27}$$

By the even and odd symmetry, it can be shown that

$$G_{R+} = -G_{R-} = G_R$$

$$G_{I+} = G_{I-} = G_I.$$

Applying all of this to (25) gives

$$\begin{aligned} U_t(\tau) &= -2j(1/T)\{e^{j2\pi\tau/T}(G_{R+} + jG_{I+}) \\ &\quad + e^{-j2\pi\tau/T}(G_{R-} + G_{I-})\} \\ &= (4/T)\{G_R \sin 2\pi\tau/T - G_I \cos 2\pi\tau/T\} \\ &= -(4/T)(G_I^2 + G_R^2)^{1/2} \sin (2\pi\tau/T - \psi) \end{aligned} \tag{28}$$

where $\psi = \tan^{-1}(G_I/G_R)$ is a timing shift (in radians) caused by departure of $g(t)$ from even symmetry. Observe that the error characteristic has a sinusoidal shape: a consequence of the band limiting.

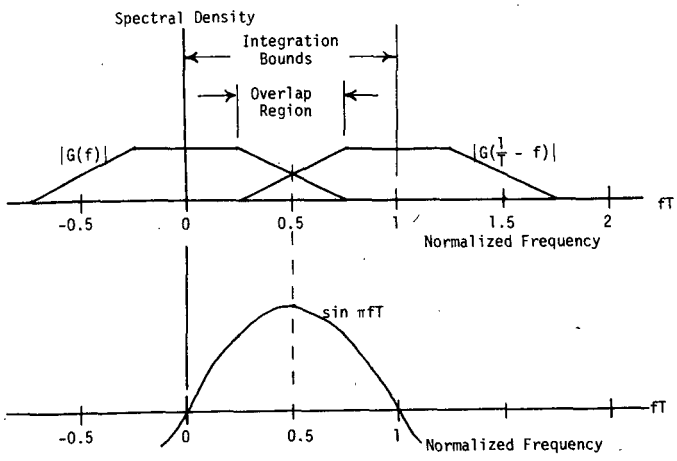


Fig. 3. Diagram of spectral overlap.

Consider a special case in which $g(t)$ is even. Optimum data filtering will match the data filter to the incoming pulse waveshape and that will always result in an even $g(t)$. If $g(t)$ is not even, the receiver is not matched to the incoming pulse. In many receivers, the filter will be at least approximately matched and so $g(t)$ will be nearly even.

If $g(t)$ is even, then $G_0(f) = 0$ for all f, so $G_I(f) = 0$ also. Furthermore, $\psi = 0$ and the average detector output reduces to

$$\begin{aligned} U_t(\tau) &= -(4/T)G_R \sin 2\pi\tau/T \\ &= -(4/T) \sin 2\pi\tau/T \int_0^{1/T} G(f) \\ &\quad \cdot G\left(\frac{1}{T} - f\right) \sin\ \pi f T\ df. \end{aligned} \tag{29}$$

The average vanishes at $\tau = 0$, corresponding to the center of each pulse, where the eye opening is maximum.

In (29), the integrand is proportional to the product $G(f)G((1/T) - f)$, which is sketched in Fig. 3. Only the region of overlap between these two functions contributes to protoclock generation. As excess bandwidth decreases, the overlap region shrinks and the gain of the detector becomes very small. Note how sin $\pi f T$ (induced by the delay differencing) weights the overlap region preferentially.

Appendix B

Effect of Carrier Phase Error

It is desirable that the timing loop be capable of acquiring lock without depending upon prior lock of the carrier loop. In this section we show that the proposed algorithm of (1) provides the same timing-error information irrespective of carrier phase. Therefore, timing can be established prior to carrier phase lock.

Represent the fictitious, time-continuous, complex signal out of the data filters as

$$w(t) = \{a(t) + jb(t)\} e^{j\Delta\theta} \tag{30}$$

which has the rectangular components

$$x_1(t) = a(t) \cos \Delta\theta - b(t) \sin \Delta\theta$$

$$x_2(t) = a(t)\ \sin\ \Delta\theta + b(t)\ \cos\ \Delta\theta \qquad (31)$$

where $\Delta\theta$ is the carrier-phase tracking error, considered arbitrary but fixed. For acquisition purposes, or for a two-dimensional QPSK signal, the timing-detector algorithm is

$$u_t(t) = x_1(t - T/2)\{x_1(t) - x_1(t - T)\} + x_2(t - T/2)\{x_2(t) - x_2(t - T)\}. \qquad (32)$$

That is, the algorithm of (11)—expressed here in the continuous-time domain—is applied to both arms of the baseband signal and the two computations are summed.

Substituting (31) into (32), and performing the appropriate trigonometry yields

$$u_t(t) = a(t - T/2)\{a(t) - a(t - T)\} + b(t - T/2)\{b(t) - b(t - T)\} \qquad (33)$$

which is quite independent of $\Delta\theta$. All terms containing $\Delta\theta$ have either cancelled, or else combined according to $\sin^2 \Delta\theta + \cos^2 \Delta\theta = 1$. The two-arm algorithm delivers the same error indication irrespective of carrier phase. Therefore, the timing loop can lock prior to locking, or even closing, of the phase loop.

References

[1] K. H. Mueller and M. Müller, "Timing recovery in digital synchronous data receivers," *IEEE Trans. Commun.*, vol. COM-14, pp. 516–530, May 1976.

[2] C.-P. J. Tzeng, D. A. Hodges, and D. G. Messerschmitt, "Timing recovery in digital subscriber loops using baudrate sampling," in *Conf. Rec. Int. Conf. Commun.*, 1985, vol. 3, paper 37.6.

[3] A. Jennings and B. R. Clarke, "Data-sequence selective timing recovery for PAM systems," *IEEE Trans. Commun.*, vol. COM-33, pp. 729–731, July 1985.

[4] M. A. Marsan and S. Benedetto, "Timing and carrier recovery for high speed data transmission over telephone channels," in *Nat. Telecommun. Conf. Rec.*, 1979, vol. 2, paper 24.5.

[5] M. A. Marsan, G. Albertengo, and S. Benedetto, "High speed modem with microprocessors: Design and implementation of the synchronization algorithms," in *Nat. Telecommun. Conf. Rec.*, 1981, vol. 1, paper B3.4.

[6] T. Suzuki, H. Takatori, M. Ogawa, and K. Tomooka, "Line equalizer for a digital subscriber loop employing switched capacitor technology," *IEEE Trans. Commun.*, vol. COM-30, pp. 2074–2082, Sept. 1982.

[7] O. Agazzi, C.-P. J. Tzeng, D. G. Messerschmitt, and D. A. Hodges, "Timing recovery in digital subscriber loops," *IEEE Trans. Commun.*, vol. COM-33, pp. 558–569, June 1985.

[8] W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*. Englewood Cliffs, NJ: Prentice-Hall, 1973, ch. 9.

[9] L. E. Franks, "Carrier and bit synchronization in data communication," *IEEE Trans. Commun.*, vol. COM-28, pp. 1107–1120, Aug. 1980.

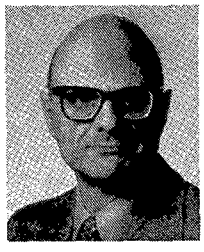
[10] F. M. Gardner, "Carrier and clock synchronization for TDMA digital communications," Euro. Space Agency, Noordwijk, the Netherlands, Rep. ESA TM-169 (ESTEC), Dec. 1976.

[11] L. E. Franks and J. P. Bubrouski, "Statistical properties of timing jitter in a PAM timing recovery scheme," *IEEE Trans. Commun.*, vol. COM-22, pp. 913–920, July 1974.

[12] U. Mengali, "A self bit synchronizer matched to the signal shape," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-7, pp. 686–693, July 1971.

[13] F. M. Gardner, "Clock and carrier synchronization: Prefilter and antihangup investigations," Euro. Space Agency, Noordwijk, the Netherlands, Rep. ESA CR-984, Contr. 2582/75, Nov. 1977.

★



Floyd M. Gardner (S'49–A'54–SM'58–F'80) received the B.S.E.E. degree from the Illinois Institute of Technology, Chicago, in 1950, the M.S.E.E. degree from Stanford University, Stanford, CA, in 1951, and the Ph.D. degree from the University of Illinois, Urbana, in 1953.

He has been an independent consulting engineer since 1960, active in the fields of communications and electronics. He is the author of the book *Phaselock Techniques* (New York: Wiley, 2nd ed., 1979).

Dr. Gardner is a Registered Professional Engineer in the State of California.