6 DUAL-INPUT DESCRIBING FUNCTION (DIDF)

6.0 INTRODUCTION

The two-sinusoid-input describing function (TSIDF) of Chap. 5 is certainly again brought to mind by the title of this chapter, as are all other describing functions which simultaneously accommodate two nonlinearity input waveforms. Because of the specific utility of the particular describing function of this class to be discussed presently, however, we reserve for it the otherwise general appellation *dual-input describing function*, and related abbreviation DIDF.

TSIDF application to nonlinear control systems is conceptually limited by requiring that only sinusoids be described in the nonlinearity test input. For this reason an alternative linearization accommodating two inputs was sought by researchers interested, among other things, in the approximate forced response behavior of certain nonlinear systems. The DIDF, as we shall now see, is a physically motivated linearization of a nonlinearity which readily permits study, among other things, of the forced responses of *limit cycling* nonlinear systems. In what follows it is required that command

inputs do not cause the limit cycle to terminate, an assumption that can be verified when under suspicion. The nonlinearity input is then comprised of the limit cycle and a component due to the command input. By assuming that the component due to the command input varies little during a limit cycle period, one can formulate a nonlinearity linearization similar in concept to the TSIDF, *but far simpler to calculate.* Hence *the DIDF model input waveform is a bias plus a sinusoid,* the latter component effectively serving to linearize the nonlinearity gain to the former.

The idea of linearizing nonlinear characteristics by means of an additive sinusoid is not a new one. MacColl (Ref. 12) described the use of such a signal in a motor-drive system incorporating a relay. Loeb (Ref. 10) has suggested that *any* nonlinear system can be treated in this manner. Lozier (Ref. 11) is credited with a method of treating oscillating control systems using this linearization, an interpretation later independently arrived at by Li and Vander Velde (Ref. 9) in connection with limit cycling adaptive feedback control system applications. Other papers by Gelb (Refs. 3–5) followed this interpretation and further developed the dynamic characterization of limit cycling systems. Oldenburger experimentally discovered the effects of an *additive* high-frequency low-amplitude input to a control system, and subsequently provided analytical justification via DIDF considerations for these effects in several papers on "signal stabilization" (Refs. 16–18). Popov (Ref. 20) and later Popov and Pal'tov (Ref. 21) have published books in which DIDF harmonic linearization is treated.

In this chapter, following general formulation and calculation of several specific DIDFs, the forced response of a class of limit cycling systems is considered. The role of poles and zeros in the description of limit cycling systems is established in this context, and application of the theory is made to a class of adaptive control systems. Adaptive roll control for a missile is considered as an example in this class. It is then shown that limit cycles in systems with an *asymmetric* nonlinearity are easily determined. The uses of artificial dither for the compensatory purposes of linearization and signal stabilization are treated. Finally, a brief TSIDF derivation is developed in DIDF terms.

Before embarking upon a discussion of DIDF formulation, we digress to examine a system quite different from those discussed in the remainder of this chapter.

Example 6.0-1 Apart from limit cycling systems, the DIDF model input waveform can arise in quite a number of different ways. For illustration, consider the problem of determining the near-circular-orbit period of an earth satellite. The coupled radial and tangential force equations are

$$\ddot{r} - r\dot{\psi}^2 = -\frac{\mu}{r^2} \tag{6.0-1}$$

and

$$r\ddot{\psi} + 2\dot{r}\dot{\psi} = 0 \tag{6.0-2}$$

or since Eq. (6.0-2) can be written as $(1/r)\, d/dt(r^2\dot{\psi}) = 0$, we can write

$$r^2\dot{\psi} = h \tag{6.0-3}$$

where r and ψ are as defined in Fig. 6.0-1, μ is a constant defining the specific gravitational force, and h is the constant specific angular momentum of the orbit. Eliminating $\dot{\psi}$ from Eq. (6.0-1) by means of Eq. (6.0-3) gives

$$\ddot{r} - \frac{h^2}{r^3} = -\frac{\mu}{r^2} \tag{6.0-4}$$

A natural approximate solution to this conservative equation proceeds as follows: First, the geometry of the problem suggests a solution of the form ($\psi = \omega t$)

$$r = R + \delta_R \cos\psi \tag{6.0-5}$$

where R and δ_R are constant. Substituting this expression in Eq. (6.0-4) gives

$$-\omega^2\delta_R \cos\psi - \frac{h^2}{R^3}\left(1 + \frac{\delta_R}{R}\cos\psi\right)^{-3} = -\frac{\mu}{R^2}\left(1 + \frac{\delta_R}{R}\cos\psi\right)^{-2} \tag{6.0-6}$$

The assumption of a near-circular orbit implies that $\delta_R/R \ll 1$. Correspondingly, the nonlinear terms in Eq. (6.0-6) can be approximated by appropriate first-order expansions, viz.,

$$-\omega^2\delta_R \cos\psi - \frac{h^2}{R^3}\left(1 - \frac{3\delta_R}{R}\cos\psi\right) \approx \frac{-\mu}{R^2}\left(1 - \frac{2\delta_R}{R}\cos\psi\right) \tag{6.0-7}$$

Balancing harmonics on each side of this equation (a concept which is integral to DF analysis of previous chapters) gives the relationships

$$R = \frac{h^2}{\mu} \qquad \text{and} \qquad \omega^2 = \frac{3h^2}{R^4} - \frac{2\mu}{R^3} \tag{6.0-8}$$

Solution of these two equations for the orbit period T finally yields

$$T = \frac{2\pi}{\omega} = 2\pi\frac{h^3}{\mu^2} \tag{6.0-9}$$

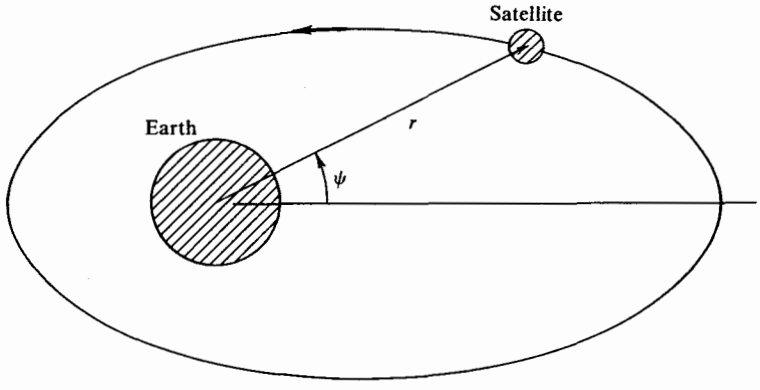


Figure 6.0-1 *Earth-satellite geometry.*

This compares favorably with the exact result (where orbit eccentricity $= \delta_R/R$)

$$T_{\text{exact}} = 2\pi \frac{h^3}{\mu^2} \left[1 - \left(\frac{\delta_R}{R} \right)^2 \right]^{-\frac{3}{2}} \tag{6.0-10}$$

The ease with which this approximate solution is obtained is characteristic of the harmonic linearization approach. This accounts, in large measure, for the popularity and widespread use of describing function techniques.

6.1 MATHEMATICAL FORMULATION OF THE DIDF

MOTIVATION

Consider the system of Fig. 6.1-1 to be in a limit cycle state of period T. Now let a "slowly varying," but otherwise arbitrary, input which satisfies the inequality

$$T \left| \frac{dr(t)}{dt} \right| \ll A \tag{6.1-1}$$

be applied, where A is the amplitude of the limit cycle oscillation at the input to N. From this inequality it can be seen that a slowly varying function is taken as one which changes relatively little with respect to A over the period T. This coarse, somewhat restrictive definition is sufficient for our present purposes.

In this case a typical set of input and output waveforms are illustrated in Fig. 6.1-2*a*, where we see that the output follows the input, on the average, to within some dynamic following error. With the exception of an additive limit cycle component this is the input-output relationship ordinarily encountered in control systems of all types, where the function of the system is simply to reproduce the input waveform at a higher power level at the output. One cycle of the associated nonlinearity input waveform is shown in Fig. 6.1-2*b*. This model is the starting point for DIDF derivation. It plays a role completely analogous to the single-sinusoid model of DF usage, or the two-sinusoid model of TSIDF usage. Let us also note that precisely the same model is arrived at in general consideration of the limit cycle behavior of a system with an asymmetric nonlinearity, for in this case the

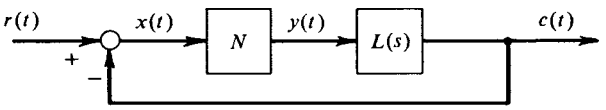


Figure 6.1-1 *Limit-cycling system.*

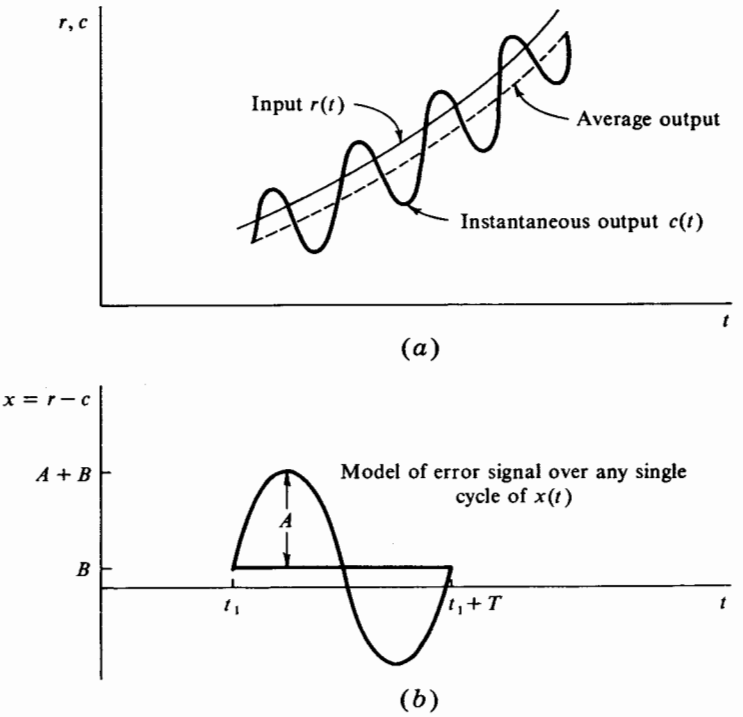


Figure 6.1-2 *(a) Typical input and output waveforms with (b) associated error-signal model.*

closed loop can develop a bias input to N in addition to the limit cycle. Other physically significant situations in which a true bias appears at the nonlinearity input, in addition to a sinusoid, include a system with an integration in the linear part responding to a ramp input while limit cycling, a limit cycling system responding to a constant disturbance input, and a non-limit-cycling system responding to a sinusoidal command and a constant disturbance input.

FORMULATION

The input to the nonlinearity is taken to be

$$x(t) = B + A \sin(\omega t + \theta) \tag{6.1-2}$$

From the statistical point of view taken in Chap. 1, B, A, and ω are considered determined by the nature of the system and its inputs. Thus the only random variable in the characterization of $x(t)$ is the phase angle θ,

which has the uniform distribution over one cycle. The expectations which appear in the expressions for the describing functions are in this case just single integrations over a 2π interval in θ.

The expressions for the describing functions resulting from the uncoupled relations of Eq. (1.5-21) are clearly applicable in this case. The nonlinearity input is the sum of two components: $x_1(t)$ may be taken to be the bias, and $x_2(t)$ the sinusoid. With B a deterministic quantity, and the sinusoid $x_2(t)$ having zero mean, these two input components are obviously uncorrelated. With the describing function for the sinusoidal input component interpreted as the sum of an in-phase plus quadrature gain, as was discussed in Chap. 1 and again in Chap. 2, the output of the optimum quasi-linear approximator to a nonlinearity with this input is

$$y_a(t) = N_B B + n_p A \sin(\omega t + \theta) + n_q A \cos(\omega t + \theta) \qquad (6.1\text{-}3)$$

The required gains, or describing functions, can readily be calculated for quite general nonlinearities. In recognition of this we indicate in the notation a dependence of nonlinearity output on the nonlinearity input $x(t)$ and its derivative.

$$y(t) = y[x(t),\dot{x}(t)] \qquad (6.1\text{-}4)$$

This, however, is not to be taken as a restriction on the form of nonlinearities which can be treated. The only requirement is that one be able to define the nonlinearity output as the phase of the input sinusoid traverses a full cycle. The describing functions will in general depend on the bias level, the amplitude of the sinusoid, and its frequency.

The approximating gain to the bias input component is given by Eq. (1.5-27).

$$\begin{aligned} N_B(A,B,\omega) &= \frac{1}{B}\overline{y(0)} \\ &= \frac{1}{2\pi B}\int_0^{2\pi} y(B + A\sin\theta,\, A\omega\cos\theta)\, d\theta \end{aligned} \qquad (6.1\text{-}5)$$

The real and imaginary parts of the complex approximating gain to the sinusoidal input component are given by Eqs. (1.5-36).

$$\begin{aligned} n_p(A,B,\omega) &= \frac{2}{A}\overline{y(0)\sin\theta} \\ &= \frac{1}{\pi A}\int_0^{2\pi} y(B + A\sin\theta,\, A\omega\cos\theta)\sin\theta\, d\theta \end{aligned} \qquad (6.1\text{-}6)$$

$$\begin{aligned} n_q(A,B,\omega) &= \frac{2}{A}\overline{y(0)\cos\theta} \\ &= \frac{1}{\pi A}\int_0^{2\pi} y(B + A\sin\theta,\, A\omega\cos\theta)\cos\theta\, d\theta \end{aligned} \qquad (6.1\text{-}7)$$

These describing functions, defined by the property of minimum mean-squared approximation error, are seen to be identical with the result of expanding the nonlinearity output, for the assumed input, into its Fourier series, and relating the corresponding input and output terms. Thus the mean output is related to the mean input by the gain N_B, and the fundamental output component is related to the sinusoidal input component by the complex gain N_A, which has real and imaginary parts n_p and n_q, respectively. This is a specialization of more general properties derived in Chap. 1. For any nonlinearity input, N_B is always the gain which equates the mean output of the quasi-linear approximator to the mean output of the nonlinearity. And for any nonlinearity input which does not include harmonically related sinusoids, N_A always represents the amplitude and phase relation between a sinusoidal input component and the harmonic component of the same frequency in the expectation of the nonlinearity output, computed by averaging over all random parameters except the phase angle of that sinusoid. In this case there are no random parameters other than the phase angle of the input sinusoid; so the expectation of the nonlinearity output referred to above is just the output itself.

It is evident from this formulation that in the limit, as $B \to 0$, $n_p(A,B,\omega)$ and $n_q(A,B,\omega)$ approach $n_p(A,\omega)$ and $n_q(A,\omega)$, the DF characterization of N. Also in the limit, as $B \to 0$, the gain to the bias approaches a value independent of B, which is termed the *incremental-input describing function.* It is defined by

$$N_B(A,\omega) = \lim_{B\to 0} [N_B(A,B,\omega)]$$

$$= \lim_{B\to 0} \left[\frac{1}{2\pi B} \int_0^{2\pi} y(B + A \sin\theta, A\omega \cos\theta)\, d\theta \right] \qquad (6.1\text{-}8)$$

Intuition leads one to suspect that this form of the incremental-input describing function is equivalent to that presented in Sec. 5.5 for non-harmonically related input sinusoids (i.e., asynchronous case). This equivalence is easily demonstrated.

Example 6.1-1 *Proof of the equality* $N_i(A) = N_B(A,0)$. Application of L'Hospital's rule to the indeterminate form of the incremental-input describing function [Eq. (6.1-8)], specialized to static single-valued nonlinearities, yields

$$N_B(A,0) = \frac{1}{2\pi} \lim_{B\to 0} \left[\frac{1}{B} \int_0^{2\pi} y(B + A \sin\theta)\, d\theta \right]$$

$$= \frac{1}{2\pi} \int_0^{2\pi} y'(A \sin\theta)\, d\theta \qquad (6.1\text{-}9)$$

where $y'(A \sin\theta)$ stands for the derivative of y with respect to its argument, computed at the point where the argument takes the value $A \sin\theta$. It is to be noted that Eq. (6.1-9)

provides a useful means for computing $N_B(A,0)$ directly, without the intermediate step of an explicit limiting process (that is, $B \to 0$). Only in the case of discontinuous nonlinear characteristics, where y' contains impulse functions, is it again necessary to consider a limiting process (cf. Prob. 6-2).

Next, $N_i(A)$ is determined in the following way: By definition

$$N_i(A) = \frac{1}{\pi A}\int_0^{2\pi} y(A\sin\theta)\sin\theta\,d\theta + \frac{A}{2}\frac{d}{dA}\left[\frac{1}{\pi A}\int_0^{2\pi} y(A\sin\theta)\sin\theta\,d\theta\right]$$

$$= \frac{1}{2\pi A}\int_0^{2\pi} y(A\sin\theta)\sin\theta\,d\theta + \frac{1}{2\pi}\int_0^{2\pi} y'(A\sin\theta)\sin^2\theta\,d\theta \qquad (6.1\text{-}10)$$

The first integral can be integrated by parts to yield

$$\int_0^{2\pi} y(A\sin\theta)\sin\theta\,d\theta = -y(A\sin\theta)\cos\theta\Big|_0^{2\pi} + \int_0^{2\pi} Ay'(A\sin\theta)\cos^2\theta\,d\theta$$

$$= A\int_0^{2\pi} y'(A\sin\theta)\cos^2\theta\,d\theta \qquad (6.1\text{-}11)$$

Inserting Eq. (6.1-11) into Eq. (6.1-10) gives

$$N_i(A) = \frac{1}{2\pi}\int_0^{2\pi} y'(A\sin\theta)\cos^2\theta\,d\theta + \frac{1}{2\pi}\int_0^{2\pi} y'(A\sin\theta)\sin^2\theta\,d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} y'(A\sin\theta)\,d\theta \qquad (6.1\text{-}12)$$

whence it is concluded that

$$N_i(A) = N_B(A,0) \qquad (6.1\text{-}13)$$

DISCUSSION

The DIDF incremental-input describing function representation of N leads to an extremely simple and very useful description of the input-output dynamics of limit cycling systems. We shall thoroughly explore this description, following several example DIDF calculations.

The approximation used in that analysis, in addition to the normal describing function approximation, is that the bias function employed in DIDF calculation is generalized to represent arbitrary functions of time within a restricted class. This class evidently can include functions which are *slowly varying;* so over any one period of the sinusoid the function appears essentially constant. But the relation just derived, Eq. (6.1-13), states that the effective gain of the nonlinearity to a small sinusoid of any frequency in the presence of another sinusoid is equal to the effective gain of the nonlinearity to a small bias in the presence of a sinusoid. This suggests the possibility that the effective gain of the nonlinearity to a small signal of arbitrary form in the presence of a larger sinusoid may be the same, equal to

the incremental-input describing function. This suggestion is given further strength by the results of Sec. 1.5, which showed that the effective gain of a static, single-valued nonlinearity to a small signal in the presence of *any* other independent input signals is the same in the limit as the small signal approaches zero, whether it be a bias, a sinusoid, or a random process. Thus we also include in the class of arbitrary functions for which the incremental-input describing function can be employed, functions which are *small.* Usually this "smallness" condition requires that the arbitrary signal be small compared with another signal at the nonlinearity input, such as the sinusoid in the DIDF formulation. In some instances, however, the smallness must be with respect to the characteristics of the nonlinearity (cf. Example 5.5-3).

6.2 DIDF CALCULATION

Because of the great similarity between DF and DIDF calculations, only a few selected DIDF calculations are presented. These are chosen from the ranks of frequency-independent nonlinearities which are piecewise-linear symmetric, piecewise-linear asymmetric, and polynomial, as well as frequency-dependent nonlinearities. First, let us examine a general calculation of some importance.

Consider a static, but otherwise arbitrary, nonlinearity $y = y(x)$. Using the obvious relationship $d(B + A \sin \psi) = A \cos \psi \, d\psi$, it follows that $n_q(A,B)$ can be written as

$$\begin{aligned} n_q(A,B) &= \frac{1}{\pi A} \int_0^{2\pi} y(B + A \sin \psi) \cos \psi \, d\psi \\ &= \frac{1}{\pi A^2} \oint y(B + A \sin \psi) \, d(B + A \sin \psi) \\ &= \frac{1}{\pi A^2} \oint y(x) \, dx \\ &= \frac{-S}{\pi A^2} \end{aligned} \tag{6.2-1}$$

(*Note:* We have returned to the dummy variable ψ to be consistent with previous chapters)

where S is the area enclosed by $y(x)$ as x varies through a complete cycle between the limits $B + A$ and $B - A$. The reader can convince himself with the aid of a simple sketch that the minus sign in Eq. (6.2-1) is indeed appropriate.

Several conclusions of importance can be drawn from this result. The most obvious is that $n_q(A,B) = 0$ for *any* memoryless nonlinearity, since here $S = 0$. Also, since $n_q(A,B)$ is equal to $n_q(A)$ when B is set equal to zero, it follows that Eq. (6.2-1) is suitable for calculating the phase-shifting component of the DF, $n_q(A)$.

Now we proceed to specific nonlinearities.

IDEAL RELAY

For this nonlinearity the input and output waveforms are illustrated in Fig. 6.2-1. One can see clearly that, because of the presence of a positive B, the points at which the nonlinearity output switches between D and $-D$ are altered in such a way that a net positive bias component results in y. From Eqs. (6.1-6) and (6.1-7), in which frequency dependence is dropped, we get for this *static* nonlinearity

$$\begin{aligned} n_p(A,B) &= \frac{1}{\pi A}\int_0^{2\pi} y(B + A\sin\psi)\sin\psi\, d\psi \\ &= \frac{1}{\pi A}\left[\int_0^{\pi+\psi_1} D\sin\psi\, d\psi + \int_{\pi+\psi_1}^{2\pi-\psi_1}(-D)\sin\psi\, d\psi \right. \\ &\qquad \left. + \int_{2\pi-\psi_1}^{2\pi} D\sin\psi\, d\psi\right] \\ &= \frac{4D}{\pi A}\cos\psi_1 \\ &= \frac{4D}{\pi A}\sqrt{1-\left(\frac{B}{A}\right)^2} \end{aligned} \tag{6.2-2}$$

and

$$\begin{aligned} n_q(A,B) &= \frac{1}{\pi A}\int_0^{2\pi} y(B + A\sin\psi)\cos\psi\, d\psi \\ &= \frac{1}{\pi A}\left[\int_0^{\pi+\psi_1} D\cos\psi\, d\psi + \int_{\pi+\psi_1}^{2\pi-\psi_1}(-D)\cos\psi\, d\psi \right. \\ &\qquad \left. + \int_{2\pi-\psi_1}^{2\pi} D\cos\psi\, d\psi\right] \\ &= 0 \end{aligned}$$

This last result is obtained with even greater ease by inspection of Eq. (6.2-1). The limit cycle DIDF is indeed non-phase-shifting, as one expects in the

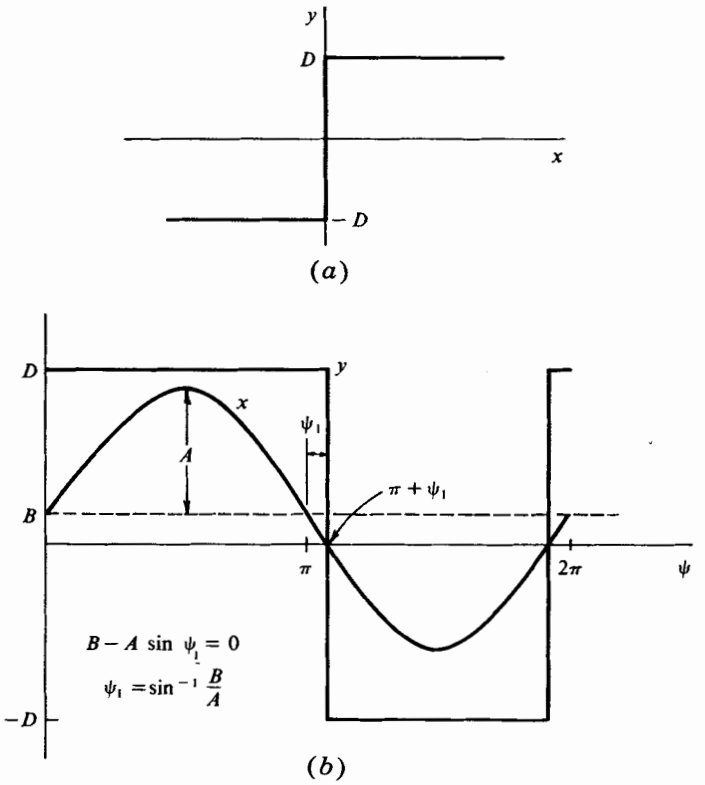


Figure 6.2-1 *(b) Input and output waveforms for the ideal relay (a).*

case of this memoryless nonlinearity. From Eq. (6.1-5) the signal DIDF is easily determined as

$$N_B(A,B) = \frac{1}{2\pi B}\int_0^{2\pi} y(B + A\sin\psi)\,d\psi$$

$$= \frac{1}{2\pi B}\left[\int_0^{\pi+\psi_1} D\,d\psi + \int_{\pi+\psi_1}^{2\pi-\psi_1}(-D)\,d\psi + \int_{2\pi-\psi_1}^{2\pi} D\,d\psi\right]$$

$$= \frac{2D}{\pi B}\psi_1$$

$$= \frac{2D}{\pi B}\sin^{-1}\frac{B}{A} \tag{6.2-3}$$

The incremental-input describing function follows directly.

$$N_B(A) = \lim_{B \to 0} N_B(A,B) = \frac{2D}{\pi A} \tag{6.2-4}$$

The ease with which these calculations have been effected is striking by comparison with the corresponding TSIDF calculations of the previous chapter. As stated earlier, it is desired to let B represent a multitude of signals in addition to a simple bias. As an illustration of the case where B represents a sinusoid, let us examine the above ideal-relay results in order to get a feel for the conditions under which the DIDF is an acceptable substitute for the TSIDF. From the TSIDF ideal-relay calculation in the general case where the signal frequency is not a rational fraction of the limit cycle frequency, it is shown in Eqs. (5.1-19) that ($A > |B|$)

$$\begin{aligned} N_A(A,B) &= \frac{4D}{\pi A}\left[1 - \frac{1}{4}\left(\frac{B}{A}\right)^2 - \frac{3}{64}\left(\frac{B}{A}\right)^4 - \cdots\right] \\ N_B(A,B) &= \frac{2D}{\pi A}\left[1 + \frac{1}{8}\left(\frac{B}{A}\right)^2 + \frac{3}{64}\left(\frac{B}{A}\right)^4 + \cdots\right] \end{aligned} \tag{6.2-5}$$

Corresponding DIDF expressions [Eqs. (6.2-2) and (6.2-3)] can be expanded to yield

$$\begin{aligned} N_A(A,B) &= \frac{4D}{\pi A}\left[1 - \frac{1}{2}\left(\frac{B}{A}\right)^2 - \frac{1}{8}\left(\frac{B}{A}\right)^4 - \cdots\right] \\ N_B(A,B) &= \frac{2D}{\pi A}\left[1 + \frac{1}{6}\left(\frac{B}{A}\right)^2 + \frac{3}{40}\left(\frac{B}{A}\right)^4 + \cdots\right] \end{aligned} \tag{6.2-6}$$

Examining TSIDF calculations for the cases of rational, as well as irrational, frequency ratios and including *all* relative phase shifts leads to the conclusion that the ideal-relay DIDF result is within 5 percent of the TSIDF result under the conditions

$$\begin{aligned} &\text{Amplitude-ratio condition:} && \frac{B}{A} < \frac{1}{3} \\ &\text{Frequency-ratio condition:} && \gamma < \frac{1}{3} \end{aligned} \tag{6.2-7}$$

These rough quantitative statements apply as well to a wide range of common nonlinearities. They are therefore adopted as guideposts in a limit cycling system input-output characterization.

At this point another interpretation of the significance of the describing function $N_B(A,B)$ can be stated. Given that there is a limit cycle of amplitude A, this quantity explicitly accounts for the transmission of slowly varying signals through the nonlinearity in the presence of the limit cycle.

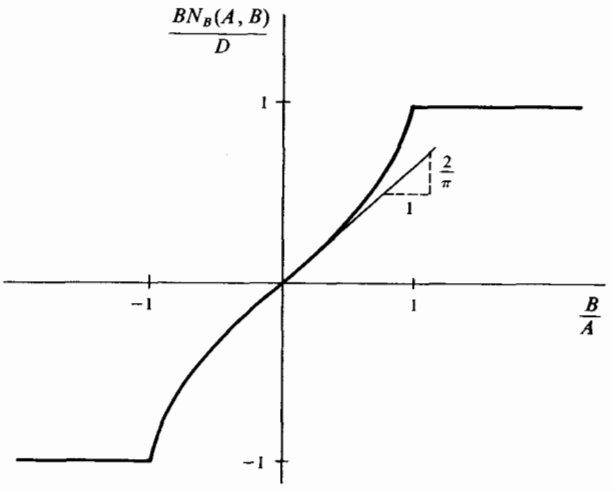


Figure 6.2-2 Equivalent nonlinear element for signal transmission through an ideal relay in the presence of a limit cycle.

It defines an *equivalent nonlinear element* (sometimes called a "modified nonlinearity") in the sense that one can divorce limit cycle considerations from consideration of nonlinearity signal transmission, provided that A is constant. A normalized plot of $N_B(A,B)$ appears in Fig. 6.2-2. Insofar as the values of B/A encountered may be small, this *equivalent nonlinear element* can be replaced by an *equivalent linear element*, given by its slope at the origin, $2D/\pi A$. This is the physical interpretation of the incremental-input describing function computed earlier.

RECTANGULAR HYSTERESIS

This piecewise-linear characteristic, possessing memory, leads to a square-wave output, as shown in Fig. 6.2-3. It is clear that the fundamental component of the square wave is *not* in phase with the sinusoidal part of x. Thus the limit cycle DIDF is sought in the compact form

$$\begin{aligned} N_A(A,B) &= \frac{j}{\pi A}\int_0^{2\pi} y(B + A\sin\psi)e^{-j\psi}\,d\psi \\ &= \frac{j}{\pi A}\left[\int_0^{\psi_1}(-D)e^{-j\psi}\,d\psi + \int_{\psi_1}^{\pi+\psi_2} De^{-j\psi}\,d\psi + \int_{\pi+\psi_2}^{2\pi}(-D)e^{-j\psi}\,d\psi\right] \\ &= \frac{2D}{\pi A}(e^{-j\psi_1} + e^{-j\psi_2}) \\ &= \frac{2D}{\pi A}(e^{-j\sin^{-1}[(\delta-B)/A]} + e^{-j\sin^{-1}[(\delta+B)/A]}) \end{aligned} \tag{6.2-8}$$

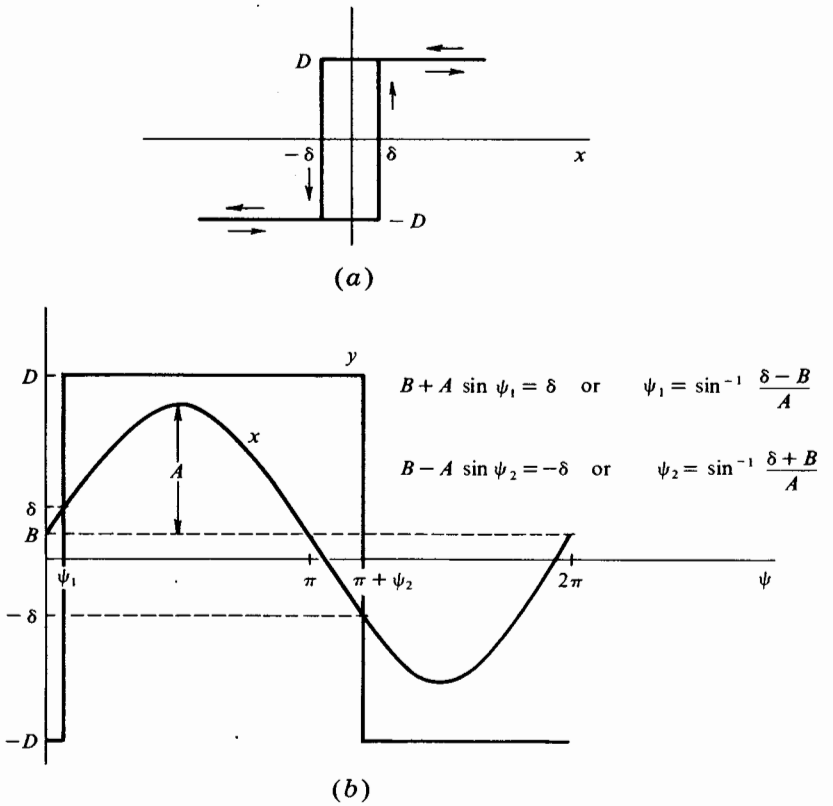


Figure 6.2-3 (*b*) *Input and output waveforms for the rectangular hysteresis nonlinearity* (*a*).

This result can, of course, be expanded to place in evidence either the real and imaginary or magnitude and phase-shift limit cycle DIDF terms. It is to be noted that, as $B \to 0$, Eq. (6.2-8) reduces to the DF derived for the rectangular hysteresis nonlinearity, Eq. (2.3-26).

The signal DIDF and incremental-input describing function follow directly.

$$\begin{aligned} N_B(A,B) &= \frac{1}{2\pi B}\int_0^{2\pi} y(B + A\sin\psi)\,d\psi \\ &= \frac{1}{2\pi B}\left[\int_0^{\psi_1}(-D)\,d\psi + \int_{\psi_1}^{\pi+\psi_2} D\,d\psi + \int_{\pi+\psi_2}^{2\pi}(-D)\,d\psi\right] \\ &= \frac{D}{\pi B}(\psi_2 - \psi_1) \\ &= \frac{D}{\pi B}\left[\sin^{-1}\left(\frac{\delta + B}{A}\right) - \sin^{-1}\left(\frac{\delta - B}{A}\right)\right] \end{aligned} \tag{6.2-9}$$

Expressing the incremental-input describing function calculation as a process of differentiation yields

$$\begin{aligned} N_B(A) &= \lim_{B\to 0} N_B(A,B) \\ &= \frac{2D}{\pi A} \lim_{B/A\to 0} \frac{\sin^{-1}(\delta/A + B/A) - \sin^{-1}(\delta/A - B/A)}{2B/A} \\ &= \frac{2D}{\pi A} \frac{d}{d(\delta/A)} \sin^{-1}\frac{\delta}{A} \\ &= \frac{2D}{\pi A} \frac{1}{\sqrt{1-(\delta/A)^2}} \end{aligned} \tag{6.2-10}$$

TWO-SEGMENT PIECEWISE-LINEAR ASYMMETRIC NONLINEARITY

This memoryless characteristic is shown in Fig. 6.2-4. By proper choice of m_1 and m_2, it can be made to represent an absolute-value device ($m_2 = -m_1$), a rectifier voltage-current characteristic ($m_1 \gg m_2 > 0$), and so forth. DIDF calculation proceeds easily. Since the nonlinearity is memoryless, it results that ($|B| \leq A$)

$$\begin{aligned} N_A(A,B) &= \frac{1}{\pi A}\int_0^{2\pi} y(B + A\sin\psi)\sin\psi\, d\psi \\ &= \frac{1}{\pi A}\left[\int_0^{\pi+\psi_1} m_1(B + A\sin\psi)\sin\psi\, d\psi \right. \\ &\qquad + \int_{\pi+\psi_1}^{2\pi-\psi_1} m_2(B + A\sin\psi)\sin\psi\, d\psi \\ &\qquad \left. + \int_{2\pi-\psi_1}^{2\pi} m_1(B + A\sin\psi)\sin\psi\, d\psi\right] \\ &= \frac{m_1+m_2}{2} + \frac{m_1-m_2}{\pi}\left(2\frac{B}{A}\cos\psi_1 + \psi_1 - \frac{\sin 2\psi_1}{2}\right) \\ &= \frac{m_1+m_2}{2} + \frac{m_1-m_2}{\pi}\left[\sin^{-1}\left(\frac{B}{A}\right) + \frac{B}{A}\sqrt{1-\left(\frac{B}{A}\right)^2}\right] \end{aligned} \tag{6.2-11}$$

The term in brackets has already been found to occur frequently in DF calculations [denoted $f(B/A)$ in Sec. 2.3]. In terms of the previously introduced notation, Eq. (6.2-11) can be written as

$$N_A(A,B) = \frac{m_1+m_2}{2} + \frac{m_1-m_2}{2} f\left(\frac{B}{A}\right)$$

It is to be observed that the above results are valid only for a restricted range of B. Outside of this range, inspection yields

$$N_A = \begin{cases} m_1 & \text{for } B > A \qquad (6.2\text{-}12) \\ m_2 & \text{for } B < -A \qquad (6.2\text{-}13) \end{cases}$$

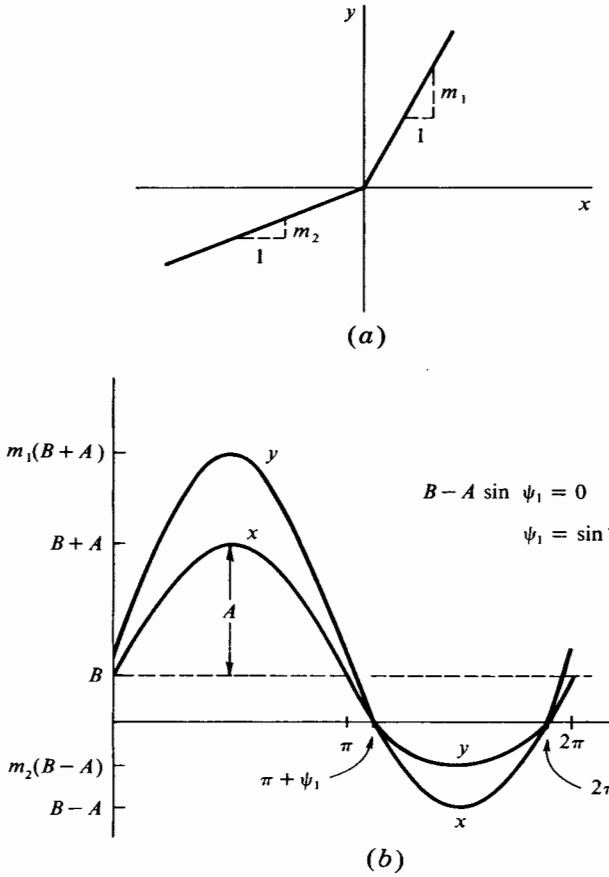


Figure 6.2-4 *(b) Input and output waveforms for a two-segment piecewise-linear asymmetric nonlinearity (a).*

Continuing,

$$N_B(A,B) = \frac{1}{2\pi B}\int_0^{2\pi} y(B + A\sin\psi)\,d\psi$$

$$= \frac{1}{2\pi B}\left[\int_0^{\pi+\psi_1} m_1(B + A\sin\psi)\,d\psi + \int_{\pi+\psi_1}^{2\pi-\psi_1} m_2(B + A\sin\psi)\,d\psi + \int_{2\pi-\psi_1}^{2\pi} m_1(B + A\sin\psi)\,d\psi\right]$$

$$= \frac{m_1 + m_2}{2} + \frac{m_1 - m_2}{\pi}\left(\psi_1 + \frac{A}{B}\cos\psi_1\right)$$

$$= \frac{m_1 + m_2}{2} + \frac{m_1 - m_2}{\pi}\frac{A}{B}\left[\frac{B}{A}\sin^{-1}\left(\frac{B}{A}\right) + \sqrt{1 - \left(\frac{B}{A}\right)^2}\right] \quad (6.2\text{-}14)$$

The term in brackets here differs from that in Eq. (6.2-11), but as in the case of $f(B/A)$, this new term occurs repeatedly in DIDF calculation. It is thus designated $(\pi/2)g(B/A)$, in which case Eq. (6.2-14) can be written as

$$N_B(A,B) = \frac{m_1 + m_2}{2} + \frac{m_1 - m_2}{2}\frac{A}{B}\,g\left(\frac{B}{A}\right)$$

Clearly, in the case of this nonlinearity, the incremental-input describing function defined by Eq. (6.1-8) is meaningless, for an output bias appears even in the absence of an input bias. A more meaningful quantity is the gain to vanishingly small perturbations about that particular input bias B^0 which results in zero output bias. B^0 satisfies the relationship

$$N_B(A,B^0) = 0 \qquad \text{for } B^0 \neq 0 \tag{6.2-15}$$

Another meaningful quantity in this instance is the perturbation in output bias caused by a perturbation about zero of the input bias.

POLYNOMIAL-TYPE NONLINEARITY

The class of nonlinearities under consideration is comprised of the odd functions

$$y(x) = c_n x^n \tag{6.2-16}$$

where n is an odd integer. The general formula for the limit cycle DIDF for a memoryless nonlinearity yields

$$\begin{aligned} N_A(A,B) &= \frac{1}{\pi A}\int_0^{2\pi} y(B + A\sin\psi)\sin\psi\,d\psi \\ &= \frac{c_n}{\pi A}\int_0^{2\pi}(B + A\sin\psi)^n \sin\psi\,d\psi \end{aligned} \tag{6.2-17}$$

Applying the binomial theorem in expansion of the integrand and integrating gives

$$\begin{aligned} N_A(A,B) &= \frac{c_n}{\pi A}\int_0^{2\pi}\left[\sum_{k=0}^{n}\frac{n!}{(n-k)!\,k!}(A\sin\psi)^{n-k}B^k\right]\sin\psi\,d\psi \\ &= \frac{c_n}{\pi A}\sum_{k=0}^{n}\frac{n!}{(n-k)!\,k!}A^{n-k}B^k\int_0^{2\pi}(\sin\psi)^{n-k}\sin\psi\,d\psi \\ &= \frac{c_n}{\pi}\sum_{k=0}^{n}\frac{n!}{(n-k)!\,k!}A^{n-k-1}B^k\int_0^{2\pi}(\sin\psi)^{n-k+1}\,d\psi \end{aligned} \tag{6.2-18}$$

Two necessary intermediate results are

For k odd:
$$\int_0^{2\pi}(\sin\psi)^{n-k+1}\,d\psi = 0 \tag{6.2-19}$$

For k even:
$$\int_0^{2\pi} (\sin\psi)^{n-k+1}\,d\psi = 4\int_0^{\pi/2} (\sin\psi)^{n-k+1}\,d\psi$$
$$= 2\sqrt{\pi}\,\frac{\Gamma\left(\frac{n-k+2}{2}\right)}{\Gamma\left(\frac{n-k+3}{2}\right)} \tag{6.2-20}$$

where $\Gamma(\lambda)$ is the gamma function of argument λ. Thus

$$N_A(A,B) = \frac{c_n}{\pi}\,2\sqrt{\pi}\sum_{k=0,2,4,\ldots}^{n} \frac{n!}{(n-k)!\,k!}\,A^{n-k-1}B^k\,\frac{\Gamma\left(\frac{n-k+2}{2}\right)}{\Gamma\left(\frac{n-k+3}{2}\right)}$$
$$= \frac{2c_n}{\sqrt{\pi}}\sum_{k(\text{even})=0}^{n-1} \frac{n!}{(n-k)!\,k!}\,A^{n-k-1}B^k\,\frac{\Gamma\left(\frac{n-k+2}{2}\right)}{\Gamma\left(\frac{n-k+3}{2}\right)} \tag{6.2-21}$$

The signal DIDF is computed as follows:

$$N_B(A,B) = \frac{1}{2\pi B}\int_0^{2\pi} y(B + A\sin\psi)\,d\psi$$
$$= \frac{c_n}{2\pi B}\int_0^{2\pi} (B + A\sin\psi)^n\,d\psi$$
$$= \frac{c_n}{2\pi B}\int_0^{2\pi}\left[\sum_{k=0}^{n} \frac{n!}{(n-k)!\,k!}\,(A\sin\psi)^{n-k}B^k\right] d\psi$$
$$= \frac{c_n}{2\pi B}\sum_{k=0}^{n} \frac{n!}{(n-k)!\,k!}\,A^{n-k}B^k\int_0^{2\pi} (\sin\psi)^{n-k}\,d\psi \tag{6.2-22}$$

The integral in the summation contributes only for k odd; hence

$$N_B(A,B) = \frac{c_n}{2\pi}\,2\sqrt{\pi}\sum_{k=1,3,5,\ldots}^{n} \frac{n!}{(n-k)!\,k!}\,A^{n-k}B^{k-1}\,\frac{\Gamma\left(\frac{n-k+1}{2}\right)}{\Gamma\left(\frac{n-k+2}{2}\right)}$$
$$= \frac{c_n}{\sqrt{\pi}}\sum_{k(\text{odd})=1}^{n} \frac{n!}{(n-k)!\,k!}\,A^{n-k}B^{k-1}\,\frac{\Gamma\left(\frac{n-k+1}{2}\right)}{\Gamma\left(\frac{n-k+2}{2}\right)} \tag{6.2-23}$$

Examining the limit of this expression as $B \to 0$ enables identification of the incremental-input describing function. The summation is first expanded, yielding one term in $B^0 = 1$ and $(n-1)/2$ other terms which disappear in the limit as $B \to 0$. Thus

$$N_B(A) = \lim_{B \to 0} [N_B(A,B)]$$
$$= \frac{nc_n}{\sqrt{\pi}} A^{n-1} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left[\frac{(n+1)}{2}\right]} \tag{6.2-24}$$

A very common odd nonlinearity is of the form

$$y = x^3 \tag{6.2-25}$$

Using the above results enables finding the DIDFs for the limit cycle and signal as

$$N_A(A,B) = \tfrac{3}{4}A^2 + 3B^2 \tag{6.2-26}$$
$$N_B(A,B) = \tfrac{3}{2}A^2 + B^2 \tag{6.2-27}$$

NONLINEAR CLEGG INTEGRATOR

Discussion of this dual-mode nonlinear integrator can be found in Chap. 2 (see Fig. 2.4-3). Figure 6.2-5 shows the input and discontinuous output over one complete cycle, $-\psi_1 \leq \psi \leq 2\pi - \psi_1$. The output of the Clegg integrator is determined in two pieces: First, by integrating the input waveform from $-\psi_1$, the point at which the input turns positive, to the literal variable ψ (for $-\psi_1 \leq \psi < \pi + \psi_1$):

$$y(B + A\sin\psi, A\omega\cos\psi) = \int_{-\psi_1/\omega}^{\psi/\omega} (B + A\sin\psi)\, d\left(\frac{\psi}{\omega}\right)$$
$$= \frac{1}{\omega}\int_{-\psi_1}^{\psi} (B + A\sin\psi)\, d\psi$$
$$= \frac{1}{\omega}[B(\psi + \psi_1) + A(\cos\psi_1 - \cos\psi)] \tag{6.2-28}$$

and second, by integrating *with zero initial conditions* from $\pi + \psi_1$, the point at which the input turns negative, to ψ (for $\pi + \psi_1 \leq \psi < 2\pi - \psi_1$):

$$y(B + A\sin\psi, A\omega\cos\psi) = \frac{1}{\omega}\int_{\pi+\psi_1}^{\psi} (B + A\sin\psi)\, d\psi$$
$$= \frac{1}{\omega}[B(\psi - \psi_1 - \pi) - A(\cos\psi_1 + \cos\psi)] \tag{6.2-29}$$

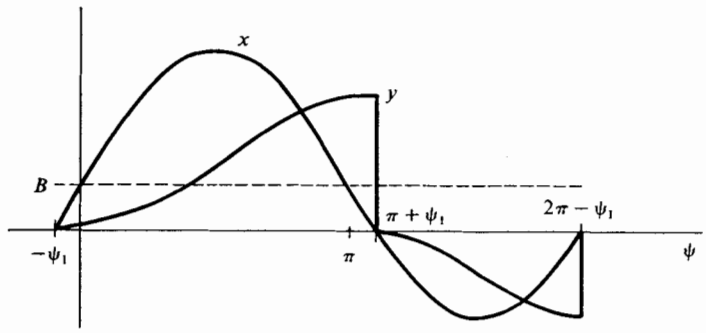


Figure 6.2-5 *Input and output waveforms for the nonlinear Clegg integrator.*

The frequency-dependent limit cycle DIDF is computed as follows:

$$\begin{aligned} N_A(A,B,\omega) &= \frac{j}{\pi A}\int_{-\psi_1}^{2\pi-\psi_1} y(B + A\sin\psi, A\omega\cos\psi)e^{-j\psi}\,d\psi \\ &= \frac{j}{\pi A\omega}\int_{-\psi_1}^{\pi+\psi_1}[B(\psi+\psi_1) + A(\cos\psi_1 - \cos\psi)]e^{-j\psi}\,d\psi \\ &\quad + \frac{j}{\pi A\omega}\int_{\pi+\psi_1}^{2\pi-\psi_1}[B(\psi-\psi_1-\pi) - A(\cos\psi_1 + \cos\psi)]e^{-j\psi}\,d\psi \\ &= \frac{4}{\pi\omega}\left[1 + \left(\frac{B}{A}\right)\sin^{-1}\left(\frac{B}{A}\right)\sqrt{1-\left(\frac{B}{A}\right)^2} - \left(\frac{B}{A}\right)^2 \right. \\ &\qquad \left. - j\frac{\pi}{4}\left(1 + 2\left(\frac{B}{A}\right)^2\right)\right] \end{aligned} \tag{6.2-30}$$

where the interval over which the DIDF is evaluated is chosen, for convenience, as $-\psi_1 \le \psi < 2\pi - \psi_1$, instead of $0 \le \psi < 2\pi$, and the relationship $\psi_1 = \sin^{-1}(B/A)$ is employed. Observe that, as it should in the limit as $B \to 0$, the limit cycle DIDF reduces to the DF computed in Chap. 2:

$$N_A(A, B = 0, \omega) = \frac{4}{\pi\omega}\left(1 \quad j\frac{\pi}{4}\right)$$

Following the same procedure gives the signal DIDF and incremental-input describing function as

$$\begin{aligned} N_B(A,B,\omega) &= \frac{1}{2\pi B}\int_{-\psi_1}^{2\pi-\psi_1} y(B + A\sin\psi, A\omega\cos\psi)\,d\psi \\ &= \frac{2}{\pi\omega}\left[\left(\frac{\pi}{2}\right)^2 + \left[\sin^{-1}\left(\frac{B}{A}\right)\right]^2 + \left(\frac{A}{B}\right)\sin^{-1}\left(\frac{B}{A}\right)\sqrt{1-\left(\frac{B}{A}\right)^2}\right] \end{aligned} \tag{6.2-31}$$

and

$$N_B(A,\omega) = \lim_{B\to 0} N_B(A,B,\omega)$$

$$= \frac{1}{\omega}\left(\frac{4+\pi^2}{2\pi}\right)$$

$$= \frac{2.21}{\omega} \tag{6.2-32}$$

Thus, in the region of small B/A, both the limit cycle and signal DIDFs are not only independent of B but of A as well. Only frequency exists as a DIDF parameter. This is certainly reminiscent of the behavior of a linear integrator. The remaining differences between these *linearized* nonlinear integrator transfer functions and the single transfer function of a linear integrator are what makes the Clegg integrator a particularly useful compensatory device.

Additional DIDF calculations are tabulated in Appendix C. The frequent appearance there of the functions $f(B/A)$ and $g(B/A)$ attest to their value as a shorthand notation.

6.3 FORCED RESPONSE OF LIMIT CYCLING NONLINEAR SYSTEMS

In this section an input-output model which accounts for transient as well as frequency response behavior is developed for limit cycling systems. Needless to say, the approximate DIDF analysis employed results in certain restrictions on the use of this model. Special attention is therefore devoted to the question of its range of validity. The results sought are approximations, such as may be of convenient use in analysis and design work. The arguments presented are both heuristic and abbreviated.

Basically, what we should like to argue is the equivalence of systems *a* and *b* of Fig. 6.3-1. That is, the original nonlinearity is to be modeled by its incremental-input describing function, and the remaining effect of the limit cycle (in addition to its modulating effect on the original nonlinear element) is to appear in the linearized system as an additive output term. The reason for employing the incremental-input describing function rather than the signal DIDF is, of course, that the equivalent system thereby becomes totally linear. A prefilter has been associated with the system to allow for a reshaping of the input amplitude spectrum such that the DIDF nonlinearity characterization is valid over the range of inputs anticipated.

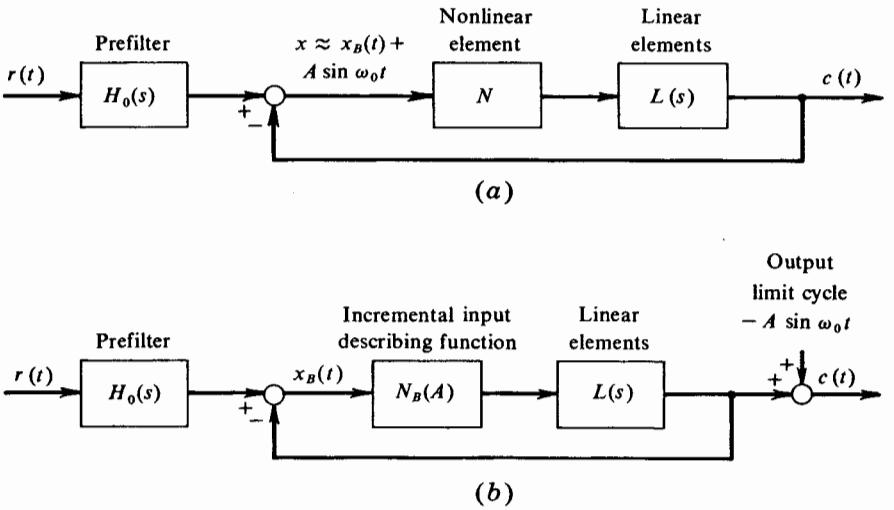


Figure 6.3-1 (*a*) *A limit cycling nonlinear system.* (*b*) *Its equivalent-linear-system model.*

POLE-ZERO SYSTEM CHARACTERIZATION

In the following discussion it is assumed that the nonlinearity input signal consists of a sinusoidal component due to the limit cycle plus another component due to the input signal (Fig. 6.3-1*a*). That is,

$$\begin{aligned} x(t) &\approx x_B(t) + x_A(t) \\ &\approx x_B(t) + A \sin \omega_0 t \end{aligned}$$

It is our intent here to discuss the meaning of poles and zeros as applied to the linear system model in Fig. 6.3-1*b*. A static nonlinearity is assumed. Consider the imaginary and real axes separately.

Along the imaginary axis we are concerned with sinusoidal response characteristics. Hence we consider the nonlinearity input signal to consist of two sinusoids, one due to the limit cycle and the other to system response to the input signal. This results in precisely the TSIDF situation studied earlier. According to the TSIDF analysis of Sec. 5.1, the gain to each sinusoid is frequency-independent provided that the sinusoidal frequencies are nonharmonically related. Since for present purposes any deterministic linking of these frequencies is not envisioned, the assumption of an irrational frequency ratio is not at all restrictive. If, further, the amplitude-ratio condition of Eq. (6.2-7) is imposed, the result for many nonlinearities is that the limit cycle amplitude is independent of the forcing signal [$N_A(A,B) \rightarrow N_A(A)$], and the gain to the smaller sinusoid is independent of its own

amplitude $[N_B(A,B) \rightarrow N_B(A)]$. Under these circumstances the system input-output description is indeed linear.

Let us now turn our attention to the axis of real exponentials. Again we assume satisfaction of the amplitude-ratio condition just cited, where the ratio now refers to peak exponential amplitude divided by peak limit cycle amplitude. In considering the nonlinearity gain to an exponential in the presence of a sinusoid, it is immediately apparent that the time duration of the exponential relative to a limit cycle period is a significant factor. Exponentials of long duration such as 10 or more limit cycle periods are certainly well represented in the DIDF input signal model consisting of sinusoid plus bias. Equation (6.1-1) tends to be satisfied in this instance. On the other hand, exponentials of sufficiently short duration can take place during various phases of the sinusoid, and the responses would be quite different. For example, in the case of an ideal relay, an exponential occurring during the limit cycle amplitude peaking would evoke essentially no additional nonlinearity output. The "gain" to such a transient signal is near zero. If, on the other hand, the same short-duration exponential occurs near a limit cycle zero crossing, the nonlinearity output indeed reflects its presence, and thus leads to a substantially larger "gain." The important fact, however, is the possible *time dependence* of nonlinearity gain to the transient signal. *As the exponential duration increases, the time dependence of this gain decreases.* Figure 6.3-2 illustrates one particular situation. Calling $\tau_{\min}$ the minimum acceptable value of exponential time constant (i.e., the time constant corresponding to maximum allowable dependence of the signal gain upon time), one could argue that the exponential should continue for a minimum of two limit cycle periods, viz.,

$$4\tau_{\min} \approx 2\frac{2\pi}{\omega_0}$$

or

$$\frac{1}{\tau_{\min}} \approx \frac{\omega_0}{3} \tag{6.3-1}$$

In this event the maximum delay (T_d in Fig. 6.3-2) is approximately 25 percent of the total exponential time duration. This value is somewhat arbitrary, but its implications will be fully apparent in the development to follow. With this order-of-magnitude calculation we proceed directly to an interpretation of the significance of poles and zeros in the complex s plane input-output system description.

Consider the s plane as divided into the three regions shown in Fig. 6.3-3. From the previous heuristic development we argue by extension that region I, to the left of the line defined by $\sigma = -\omega_0/3$, is the space in which closed-loop system poles display residues which depend upon the time they are excited. Region III, the right half-plane, cannot contain any closed-loop poles

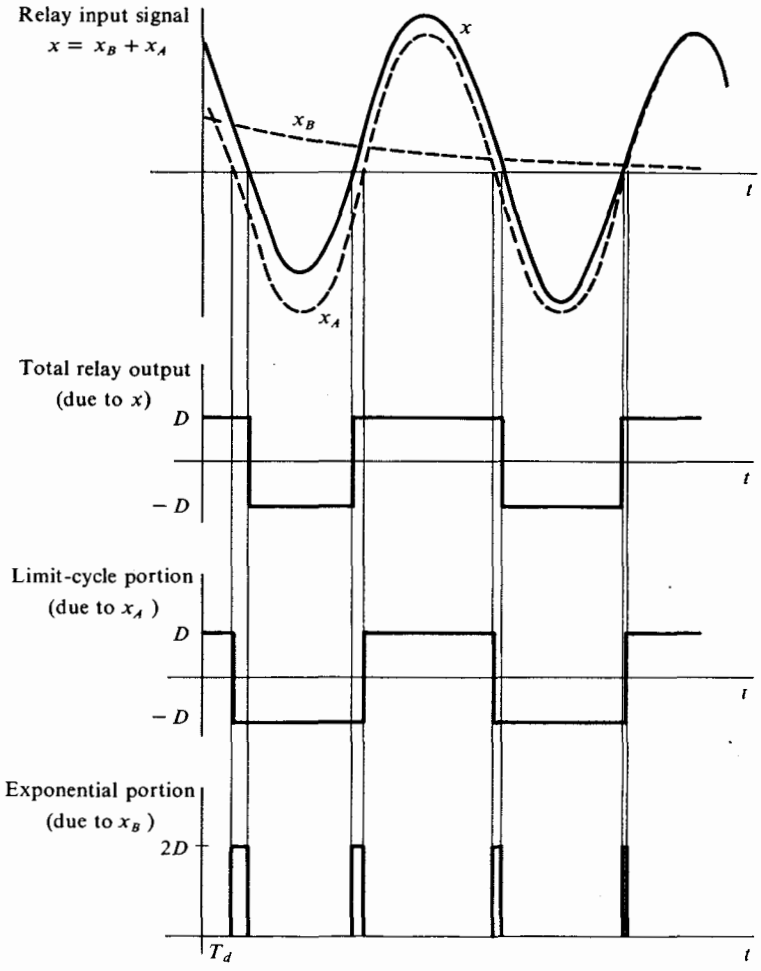


Figure 6.3-2 *Response of an ideal relay to an exponential signal in the presence of a limit cycle.*

since they would eventually violate the amplitude-ratio condition. The remaining region, designated region II, is the space wherein closed-loop poles are taken to correspond to approximately *linear time-invariant* response modes.

The addition of a zero $(1 + \tau s)$ to a function $F(s)$ yields, on a linear basis, a time function given by $f(t)$ plus $\tau\, df(t)/dt$, where $f(t) = \mathscr{L}^{-1}[F(s)]$. Since the time derivative of a sinusoid is another sinusoid of the same frequency, and since the time derivative of an exponential is another exponential with the same time constant, it is clear that the presence of zeros in no

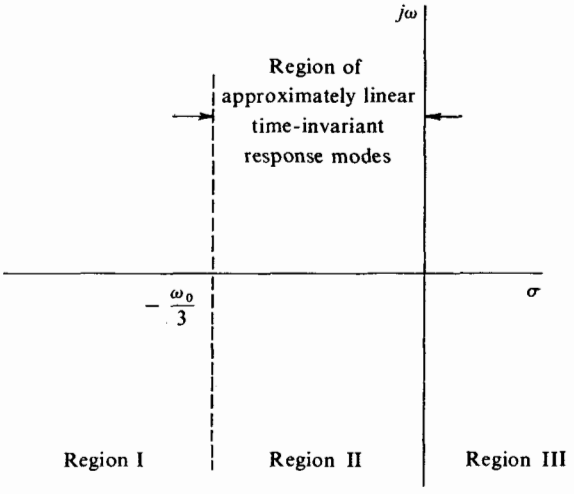
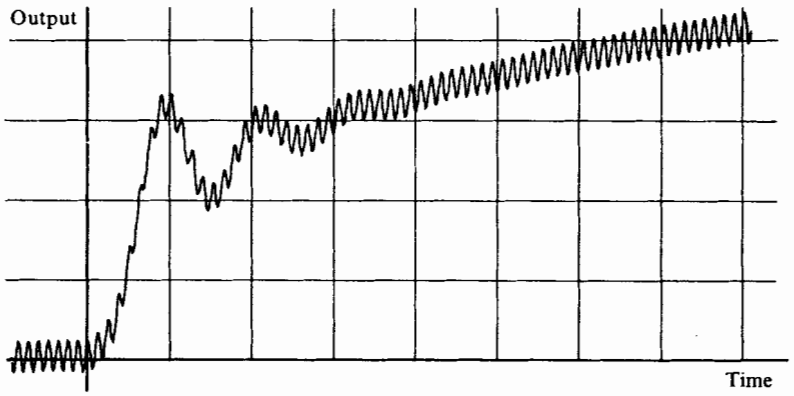


Figure 6.3-3 Regional division of the s plane.

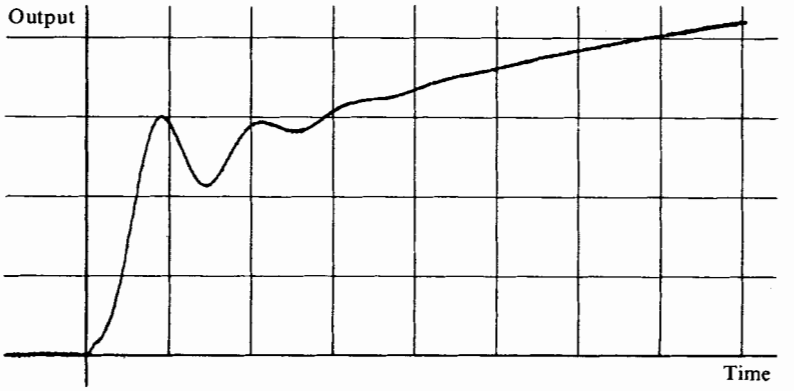
way alters the regional division of the s plane as given by Fig. 6.3-3. In fact, zeros may merely be thought of as altering the residues which accrue to the poles within the s plane. Note that all zeros must be used, regardless of the regional division in which they lie. These must be dealt with, therefore, both to secure a desired system time response and to ensure continuous satisfaction of the amplitude-ratio condition.

A limit cycling control system for a high-performance aircraft was simulated on an analog computer. The nonlinearity in this control system was an ideal relay. In order to test the equivalent gain concept for the relay in the presence of a limit cycle, transient responses for the limit cycling system and for the equivalent linear system in which the relay was replaced with a linear gain of magnitude $k = 2D/\pi A$ were studied. All predominant response modes were determined analytically to be within region II in the s plane, which thus requires an identical limit cycling system and equivalent-linear-system performance. Experimental results are shown in Fig. 6.3-4, in which the ordinate and abscissa scales for both responses are identical. Such results demonstrate the validity of the extension of pole-zero concepts to limit cycling systems, and tend to substantiate the DIDF model.

With a knowledge of the physical significance of poles and zeros for limit cycling systems, we are in a position to exploit the root-locus method, which has proved so useful in conventional linear theory. Recalling the closed-loop portion of the system of Fig. 6.3-1*b* and considering N to be a variable gain element enables construction of the locus of roots as in linear servo theory. The locus shape so derived is valid only insofar as N_B is non-phase-shifting. For $\sigma < -\omega_0/3$ sec^{-1} there is some phase shift added to signals



(*a*) Limit cycling control system



(*b*) Equivalent linear control system

Figure 6.3-4 *Transient responses of a particular limit cycling control system and its analytically computed linear equivalent.*

passing through the nonlinearity, and the actual root-locus position is uncertain to this extent. Locus gain calibration, in the usual sense, is also no longer meaningful; that is, by varying some process or compensation gain factor, the system pole locations cannot be changed. Instead, the open-loop signal gain automatically adjusts to a constant value, related to that value which establishes the limit cycle. This point will be further explored later on.

Example 6.3-1 Find the DIDF linearized equivalent system corresponding to the relay control system of Fig. 6.3-5. Assume that the prefilter has been chosen such that for all expected $r(t)$ the amplitude-ratio condition at $x(t)$ is satisfied.

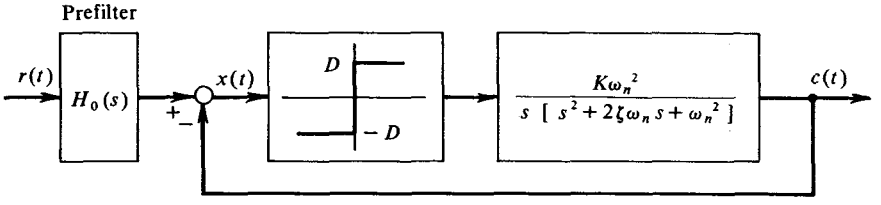


Figure 6.3-5 Third-order relay control system.

The uncalibrated root locus of the limit cycling loop is derived by assuming the relay to act as a non-phase-shifting gain, with the result shown in Fig. 6.3-6. The limit cycle frequency is determined as the point at which the locus crosses the $j\omega$ axis. Alternatively, by the DF methods of Chap. 3, the limit cycle amplitude and frequency are determined from the characteristic equation

$$s^3 + 2\zeta\omega_n s^2 + \omega_n^2 s + N_A K\omega_n^2 = 0$$

to be (Table 3.1-1)

$$\omega_n = \omega_0$$

and

$$A = \frac{2DK}{\pi\zeta\omega_n}$$

The limit cycle DIDF has been taken as

$$N_A \approx \frac{4D}{\pi A}$$

which is the value of $N_A(A,B,\omega)$ determined by Eq. (6.2-2), valid to 5 percent for all $B/A < \frac{1}{3}$. Using the above-determined value of A and the signal DIDF (incremental-input describing function),

$$N_B \approx \frac{2D}{\pi A}$$

computed from Eq. (6.2-3) by again using the fact that $B/A < \frac{1}{3}$, yields the characteristic equation of the linearized limit cycling loop as

$$s^3 + 2\zeta\omega_n s^2 + \omega_n^2 s + \zeta\omega_n^3 = 0$$

This equation has one real and two complex-conjugate roots, denoted by the squares on the three root-locus branches in Fig. 6.3-6. For small ζ, the three roots are approximately located within region II at the positions

$$s_{1,2,3} \approx -\zeta\omega_n, -\frac{\zeta\omega_n}{2} \pm j\omega_n\sqrt{1-\frac{\zeta^2}{4}}$$

with the result that

$$\frac{C}{R}(s) \approx H_0(s)\frac{\zeta\omega_n^3}{(s+\zeta\omega_n)(s^2+\zeta\omega_n s+\omega_n^2)}$$

where the prefilter transfer function has been included. Note that the gain constants K and D do not appear in the input-output transfer function. For this reason the input-output dynamics do not change with changing K and D; the system is *adaptive* with respect to these parameters. This point is more fully explored in the following sections.

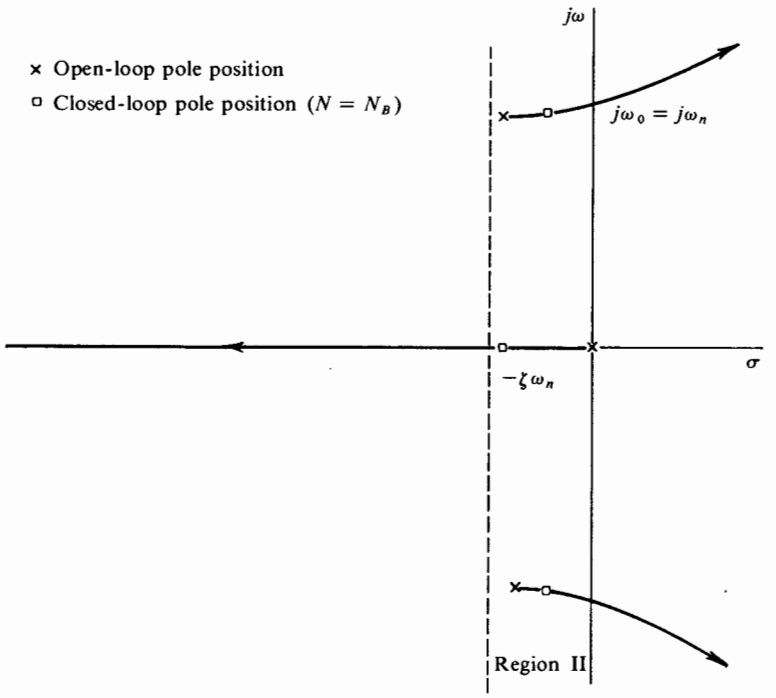


Figure 6.3-6 Root-locus diagram for the closed-loop part of the system in Fig. 6.3-5.

STABILITY

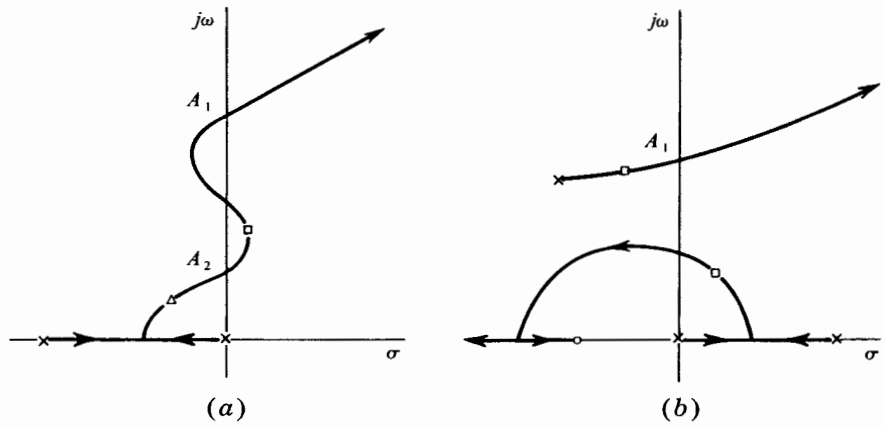
In the design of a closed-loop system for any long-term regulatory action, the first and most important specification is that the system possess no unstable modes. Linear systems with greater than 180° of open-loop phase lag, or stated differently, with root-locus branches in the right half-plane, are unstable for certain choices of open-loop gain.

On the other hand, limit cycling systems may not display similar linear unstable modes, whatever the open-loop gain setting. Consider, for example, a closed-loop control system containing an ideal relay (cf. Example 6.3-1). As the process static sensitivity K increases, the limit cycle amplitude at the nonlinearity input increases proportionally. Thus the gain N_B, inversely proportional to A, decreases, and the product KN_B is automatically held constant. This example evidences that, whereas in a linear system eventually an instability would occur, in the corresponding limit cycling system, the closed-loop dynamics do not even vary. The only price paid is the increase in limit cycle amplitude; and for cases in which this may be troublesome, Sec. 4.4 outlines the basis for its automatic regulation.

For the majority of common nonlinearities (i.e., those displaying saturation), N_B is less than N_A for $B < A$. In such cases unstable closed-loop

modes can occur only if the system is of a form which can become unstable for a decrease of open-loop gain from that value which sustains the limit cycle. The root loci for two systems of this class are illustrated in Fig. 6.3-7. In the first case, Fig. 6.3-7*a*, a conditionally stable system is illustrated. The quasi-static stability theory of Chap. 3 determines that of the three root-locus $j\omega$-axis crossings, two correspond to stable limit cycles. These are labeled A_1 and A_2, where $A_2 > A_1$. Let us assume the system to be in limit cycle state A_1. Correspondingly, the dynamics of small-signal propagation through the loop are determined by the position of all closed-loop roots associated with the nonlinearity gain $N_B(A_1)$. Since the arrows along the locus correspond to the direction of increasing open-loop gain, it follows from the relationship $N_B(A_1) < N_A(A_1)$ that the root of interest can lie on either side of the $j\omega$ axis, depending upon the gain calibration of the locus. For the situation depicted in Fig. 6.3-7*a*, this root falls in the right half-plane, indicating unstable closed-loop dynamics. This leads immediately to the conclusion that limit cycle state A_1 is actually *unstable;* only the larger-amplitude limit cycle can occur! Observe that a similar argument proves the system stable to small signals in the presence of the larger-amplitude limit cycle.

Arguing along similar lines for the unstable open-loop system of Fig. 6.3-7*b* yields that there may be no stable state whatever, depending upon the position of the lower branch root at the nonlinearity gain $N_B(A_1)$. In the



× Open-loop pole

o Open-loop zero

□ Closed-loop root corresponding to $N_B(A_1)$

△ Closed-loop root corresponding to $N_B(A_2)$

Figure 6.3-7 *Root-locus plots for a conditionally stable system (a) and an unstable open-loop system (b), both containing a static memoryless nonlinearity.*

illustration it is depicted as an unstable mode. However, if a stable situation does arise, it *must* do so at the smaller limit cycle amplitude A_1.

Let us note in passing that limit cycling control *can* be applied successfully to open-loop unstable processes. It is essential in this case to design the system not only for a stable limit cycle but also for stable closed-loop modes at the gain prescribed by N_B. In a laboratory at the Massachusetts Institute of Technology, a limit cycling system was constructed to control an inverted pendulum, with quite successful operation.

STEADY-STATE FORCED ERRORS

In response to *harmonic* forcing the steady-state forced errors are determined directly from the DIDF linearized equivalent system. To the extent that the complete representation of $N_B(A,B)$ is used in such analyses, rather than just its slope at the origin, all results obtained (in the case of unrelated frequencies) will be *identical* with TSIDF results. Section 6.8 provides the justification for this statement. It is of interest, in addition, to note the ease with which steady-state errors resulting from *aperiodic* inputs may be determined. We demonstrate by example.

Example 6.3-2 Find the steady-state following error produced by the relay system of Fig. 6.3-5 when the input is (*a*) a step of magnitude R, and (*b*) a ramp of magnitude Rt. Assume a first-order prefilter with time constant τ.

(*a*) The steady-state prefilter output is a constant, R; hence the following error of the overall system is equal to the constant-input following error of the limit cycling loop. For the limit cycling loop to be in steady state, a zero-average-value relay output is required as a result of the open-loop integration in $L(s)$. This condition can be satisfied only if $B = 0$, in which case *there is no steady-state following error*. This, of course, should come as no surprise, again because of the open-loop integration in $L(s)$.

(*b*) In the case of a ramp input, the steady-state prefilter output is another ramp, $R(t - \tau)$, displaying a following error $R\tau$. In order for the limit cycling loop to be in steady state, an average relay output equal to R/K is required, since this produces an output ramp which tracks the input. Employing the exact expressions for N_B and N_A, we get

$$\frac{R}{K} = BN_B(A,B) = \frac{2D}{\pi} \sin^{-1} \frac{B}{A}$$

or

$$\frac{B}{A} = \sin \frac{\pi R}{2DK}$$

and since $\omega_0 = \omega_n$,

$$\frac{4D}{\pi A} \sqrt{1 - \left(\frac{B}{A}\right)^2} \frac{K}{2\zeta\omega_n} = 1$$

This last condition is simply the limit cycle magnitude condition. Solving the above two equations, we get for the limit cycle amplitude

$$A = \frac{2DK}{\pi\zeta\omega_n} \cos \frac{\pi R}{2DK}$$

and for the input following error

$$B = \frac{DK}{\pi\zeta\omega_n} \sin \frac{\pi R}{DK}$$

Hence the ramp following error of the overall system is

$$\text{Ramp following error} = R\tau + \frac{DK}{\pi\zeta\omega_n} \sin \frac{\pi R}{DK}$$

Clearly, the above is only valid for $DK \geq R$. It is impossible to command any faster response from the given system. If, in addition, the limit cycle amplitude is to be relatively independent of the ramp input (say, a maximum deviation of 5 percent from the unforced condition), a drive capability $(DK)_{\min}$ is required for which

$$\cos \frac{\pi R}{2(DK)_{\min}} = 0.95$$

or

$$(DK)_{\min} \approx 5R \tag{6.3-2}$$

Notice that in this example the DIDF nonlinearity input model of a bias plus a sinusoid is approximate only to the extent that higher harmonics of the limit cycle sinusoid are omitted. The bias term is a true representation of the form of signal appearing at the nonlinearity input.

A CLOSED-LOOP PROPERTY: INHERENT ADAPTIVITY

The open-loop signal transfer function OLTF $(j\omega)$, for the closed-loop portion of the basic limit cycling control-system configuration of Fig. 6.3-8, is

$$\begin{aligned} \text{OLTF}(j\omega) &= N_B H_1(j\omega) H_2(j\omega) H_3(j\omega) L(j\omega) \\ &= N_B H(j\omega) L(j\omega) \end{aligned} \tag{6.3-3}$$

where N_B is the signal gain of the nonlinearity. Presuming satisfaction of the amplitude-ratio condition enables drawing the locus for this OLTF$(j\omega)$ on the amplitude-phase plane (gain-phase plane). Here N_B can be considered as a variable gain, shifting the system OLTF$(j\omega)$ vertically.

A first look at this method of representation for systems containing an ideal relay is facilitated by recalling the results

$$N_A \approx \frac{4D}{\pi A} \quad \text{and} \quad N_B \approx \frac{2D}{\pi A} \tag{6.3-4}$$

Correspondingly,

$$\frac{N_B}{N_A} \approx \frac{1}{2} \tag{6.3-5}$$

The interesting fact[1] that the frequency locus OLTF$(j\omega)$ always passes

[1] First discovered by Lozier (Ref. 11).

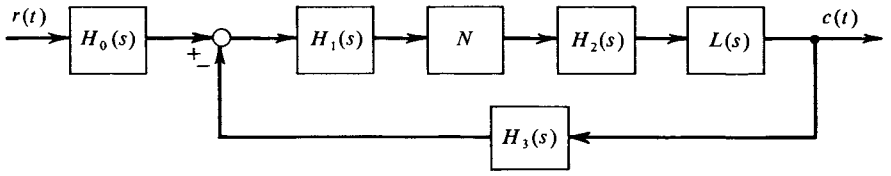


Figure 6.3-8 Basic limit cycling control-system structure.

through the point $\frac{1}{2}\angle -180°$ [or (−6 db, −180°)] on the amplitude-phase plane, to within the accuracy of Eq. (6.3-5), is easily proved as follows:

$$\begin{aligned} \text{OLTF}(j\omega) &= N_B H(j\omega) L(j\omega) \\ &= \tfrac{1}{2} N_A H(j\omega) L(j\omega) \end{aligned} \tag{6.3-6}$$

But in order to sustain a limit cycle,

$$N_A H(j\omega_0) L(j\omega_0) = -1 \tag{6.3-7}$$

therefore

$$\text{OLTF}(j\omega_0) = -\tfrac{1}{2} \tag{6.3-8}$$

The frequency locus in Fig. 6.3-9 has been drawn for a typical linear process in a limit cycling loop.

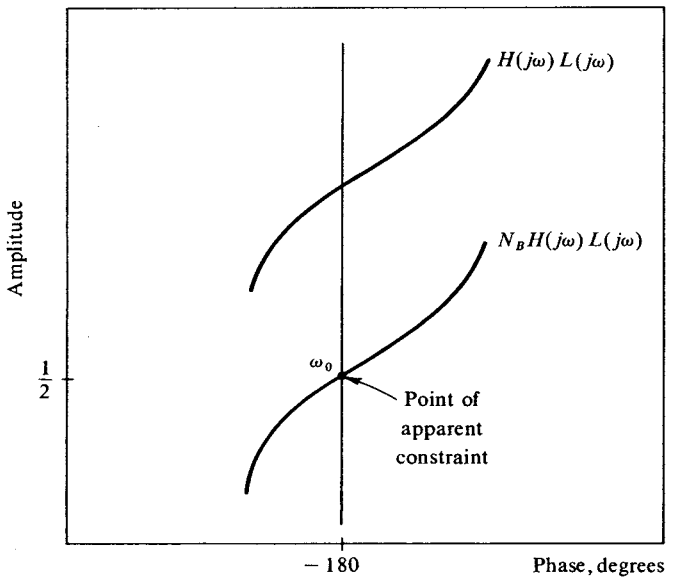


Figure 6.3-9 Typical linear elements, and associated OLTF (jω) of a relay limit cycling system containing these linear elements.

Seeing how one point of the OLTF($j\omega$) locus is constrained on the amplitude-phase plane, even in the face of changes in process characteristics, affords a first feeling for *inherent adaptivity* in a limit cycling control system. It is to be noted that Eq. (6.3-7) yields the magnitude and frequency of the limit cycle, and that for static, single-valued nonlinearities this frequency is also the frequency at which the phase of $H(j\omega)L(j\omega)$ is $-180°$. In the more general case for which the nonlinearity limit cycle DIDF is phase-shifting, the point of apparent constraint becomes

$$\text{OLTF}(j\omega_0) = -\frac{N_B}{|N_A|}\, e^{-j\angle N_A} \tag{6.3-9}$$

where the representation

$$N_A = |N_A|\, e^{j\angle N_A} \tag{6.3-10}$$

is employed.

6.4 A SCHEME FOR PARAMETER-ADAPTIVE CONTROL

Throughout this section, ξ is used to denote an "environmental parameter" such as that which causes a changing description for the controlled process $L(s,\xi)$. Thus ξ may be thought of as relating to such varied phenomena as temperature, pressure, altitude, aging, load fluctuations, and so forth. Because all physical controlled processes do in some way respond to a fluctuating environment, the need for some form of closed-loop environment insensitivity may exist. The conceptual approach to obtaining environment insensitivity beyond that provided by linear feedback configurations is termed *parameter-adaptive control.* Its objective may be stated, analytically, as the intent to continuously satisfy Eq. (6.4-1):

$$\frac{C}{R}(s,\xi) \approx \frac{C}{R}(s) \tag{6.4-1}$$

where C/R denotes the overall transfer function of the system under consideration. An approximate equality is indicated since attempts to achieve exact equality require perfect and complete *identification information*, that is, identification of the precise changes in $L(s,\xi)$. In the general case this requirement is either impractical or impossible.

In a limit cycling system the limit cycle itself is the carrier of identification information. This can be seen from Eq. (6.3-7), rewritten for convenience:

$$N_A(A)H(j\omega_0)L(j\omega_0,\xi) = -1 \tag{6.4-2}$$

The solution of this equation is

$$\text{Limit cycle at input to } N = A(\xi)\sin[\omega_0(\xi)t + \varphi] \tag{6.4-3}$$

whence we see that the limit cycle amplitude and frequency relate implicitly to the environmental parameter ξ. Although it requires a closed pool in which to circulate, the limit cycle is related *directly* to open-loop behavior. In contrast to other self-contained approaches to identification, the limit cycle adaptive control information is not attenuated by the negative feedback characteristic of the control loop. On the other hand, one limit cycle carries but two pieces of information.

Following identification is *adaptive control action*, during which the loop compensation is adjusted to maintain the limit cycle amplitude and frequency at reference values. This has the effect of constraining the amplitude and phase of the open-loop transfer function at the reference frequency, which in turn provides a good measure of environment insensitivity for the closed-loop system in many cases. Figure 6.4-1 illustrates the general limit cycling adaptive control system. The identification-data takeoff point need not be the output, and the adaptive control action feedback need not be arranged as shown. This figure is meant to be schematic only.

In this form of adaptive control there is no need to externally generate an identification test signal; the system itself provides the limit cycle. Further, there is no need for mechanizing some nonphysical performance index; the limit cycle amplitude and frequency themselves become the performance indices, and are readily available. Moreover, all instrumentation is particularly simple.

It has already been mentioned that the (ideal-relay) limit cycling control system input-output transfer is totally insensitive to controlled-process static sensitivity (zero frequency gain) perturbations. This turns out to be the case, to a first order, for many system nonlinearities besides the ideal relay.

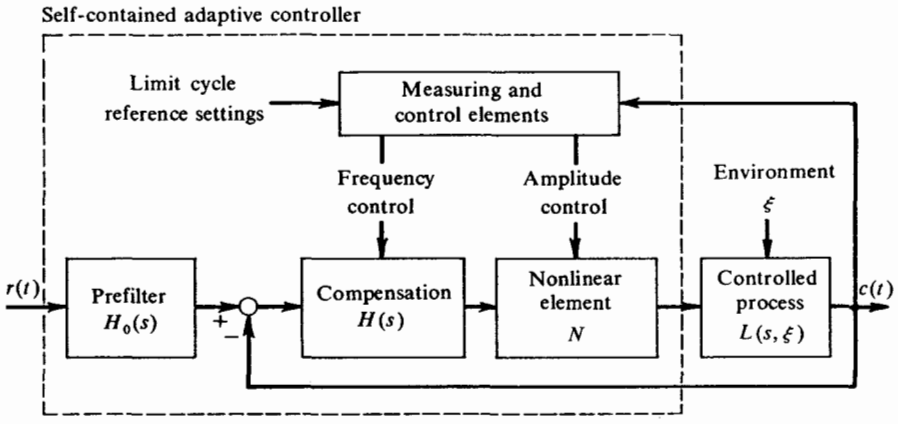


Figure 6.4-1 Limit cycling parameter-adaptive control system.

Note that this display of nearly perfect parameter adaptivity in response to controlled-process static sensitivity perturbations is a property of a limit cycling loop, the identification and adaptive control action operations occurring simultaneously and inextricably, within the nonlinearity. Controlled-process time-constant perturbations generally appear as both limit cycle amplitude and frequency perturbations, according to Eq. (6.4-2). It is the function of the measuring and control elements of Fig. 6.4-1 to sense and resolve these perturbations before directing any adaptive control action. A *sensitivity function* argument which quantitatively describes the amount of adaptivity associated with various limit cycling adaptive control schemes appears in the literature (Ref. 4).

The first commercial exploitation of a *limit cycling adaptive control system* (as such) was advanced by the Minneapolis-Honeywell Regulator Co. It is discussed in an interesting paper by Shuck (Ref. 22). The final system, an autopilot for a high-performance aircraft, was obtained by way of a long simulation program. Theoretical calculations were subsequently published (Ref. 3, pp. 294–300), which substantiated the adaptive capability of this experimental autopilot. Other applications of this technique have been determined, for instance, in connection with adaptive roll control of an air-ground missile. This application is discussed in Sec. 6.5. We digress here to consider briefly the matter of limit cycling loop compensation, both in general and with the *adaptive* control aspect in mind.

COMPENSATION

It is presumed in this chapter that the system of concern has its limit cycle mode excited at all times. This allows a simple analytic representation for the nonlinearity in the closed loop, and leads to a tractable mathematical system description. Thus a *series prefilter* is included in the limit cycling control-system structure. Its fundamental purpose is to reshape command input amplitude spectra at the input to the nonlinearity such that the amplitude-ratio condition as given by Eq. (6.2-7) is satisfied. It is clear that the prefilter design is a function of the properties of the expected class of input signals, the limit cycle, and the overall system specifications. The basic system structure is therefore as shown in Fig. 6.3-7. $H_0(s)$, $H_1(s)$, $H_2(s)$, and $H_3(s)$ denote linear filters, as is the controlled process (plant) $L(s)$. The nonlinearity N may be considered as a nonlinear filter which changes its transmittance to low-frequency input terms as the limit cycle changes. Note that it need *not* be a power element.

Compensation of a limit cycling loop requires that the nonlinearity drive level, prefilter, and loop linear filters be so chosen that system specifications are satisfied over the range of expected inputs. Thus, in the case of an ideal-relay servo, D must be chosen large enough so that the maximum values

of output derivatives which can be commanded are sufficient for following the input. In addition, if the limit cycle is to continue essentially unchanged in response to the input (a requirement for loop parameter adaptivity), the ratio of command-input-forced following error to limit cycle amplitude (B/A) must remain small. This requires a value of D which is roughly five times the input nth derivative, for a type-n plant of unity static sensitivity [Eq. (6.3-2)]. On the other hand, increasing D results in a proportional increase in A, for which a maximum specification is likely to exist at some station around the control loop.

Linear compensation is then selected to obtain the desired limit cycle frequency and amplitude. In practice, a design specification which requires a limit cycle frequency of about ten times (or greater) the desired overall system bandwidth is readily fulfilled. The position of the linear compensation is of the utmost importance. For example, if to increase the limit cycle frequency requires a *lead* network, it is most advantageously placed at H_1, because placement at H_2 or H_3 would either maximize the limit cycle amplitude at $c(t)$ or unnecessarily "slow down" the system response. Of course, there are systems in which H_1 is fixed or otherwise inaccessible. In such cases an alternative must be sought.

SUMMARY

The limit cycling adaptive control technique can be applied to all linear processes for which the existence of a continuous limit cycle is acceptable. Depending upon the nature of the variant portions of the controlled-process transfer function, limit cycling adaptive control may yield parameter-adaptive systems for which the required instrumentation is simple relative to all other current schemes. The accompanying small expense of instrumentation, coupled with comparable (and often superior) system adaptability in a number of practical situations, clearly indicates the limit cycling adaptive technique as a highly competitive approach to parameter-adaptive control. The major limitation is the compromise which usually must be struck between system drive capability (ability to follow inputs) and limit cycle amplitude. Compensation of appropriate design can often improve this compromise.

6.5 APPLICATION TO AN ADAPTIVE MISSILE ROLL CONTROL SYSTEM

One of the difficulties in designing a control system for a missile is the fact that missile aerodynamic characteristics vary greatly over the flight profile. Thus a control system which is satisfactory at one flight condition may be

completely unsatisfactory at another. One obvious solution to this problem is programmed control, in which system compensation is varied as a function of some externally measured parameters of the flight condition (such as Mach number, dynamic pressure, altitude, etc.). Generally, however, it is more desirable to have the control system "adapt" to the changing missile characteristics without the necessity of external measurements. This approach overcomes the problems of operation under uncertain environmental conditions and required a priori correlation between external measurements and actual missile characteristics, both of which can invalidate programmed control. The self-inspecting approach to design of control systems with variable controlled elements (plants) is termed *parameter-adaptive design.* In this section the limit cycling system approach to parameter-adaptive control is applied to the missile roll control problem.

In general, using this design technique, the control loop containing the variable plant dynamics is caused to limit cycle by the inclusion of a nonlinear element. In so doing, the loop thereafter automatically displays a certain insensitivity with respect to plant parameter perturbations. In the specific case at hand the nonlinearity which causes the system limit cycle is taken as part of the system fixed elements, the resulting insensitivity mentioned above providing *perfect* parameter adaptation.

HYPOTHETICAL MISSILE AND SYSTEM SPECIFICATIONS

The hypothetical vehicle to be considered is a wingless air-to-ground missile with a cruciform tail. Control moments are obtained by deflecting control surfaces on the trailing edges of the tail fins. To avert cross coupling between the vertical control surfaces and pitching motion, and between the horizontal surfaces and yawing motion, it is necessary to *roll-stabilize* the missile. Roll control moments are obtained by deflecting the control surfaces differentially.

The general equation of motion about the roll axis is

$$I_x\ddot{\varphi} + L_{\dot{\varphi}}\dot{\varphi} + L_{\varphi}\varphi = L_{\delta}\delta + L_D \tag{6.5-1}$$

where I_x = missile moment of inertia
$L_{\dot{\varphi}}$ = damping coefficient
L_{φ} = elastic restraint coefficient
L_{δ} = control torque coefficient
L_D = disturbance torques

The elastic restraint torque $L_{\varphi}\varphi$ may occur when the missile has an angle of attack and an angle of sideslip. However, it is usually so small that it may be neglected, resulting in the roll equation given in most standard texts:

$$\ddot{\varphi} + \frac{L_{\dot{\varphi}}}{I_x}\dot{\varphi} = \frac{L_{\delta}}{I_x}\delta + \frac{L_D}{I_x} \tag{6.5-2}$$

For the missile under consideration, the damping in roll is assumed to be quite small, such that the transform of Eq. (6.5-2) may be approximated by

$$\varphi(s) \approx \frac{1}{s^2}\frac{L_\delta}{I_x}\delta(s) + \frac{1}{s^2}\frac{1}{I_x}L_D(s)$$
$$= \frac{k_1(\xi)}{s^2}\delta(s) + \frac{k_2(\xi)}{s^2}L_D(s) \tag{6.5-3}$$

where s is the Laplace transform variable, and ξ is an environmental parameter. We are primarily concerned with the transfer relationship between roll angle and control-surface deflection. Since the control gain L_δ is roughly proportional to dynamic pressure, which may vary greatly over the flight profile, and in addition the moment of inertia decreases as fuel is burned, it is assumed that these variations cause $k_1(\xi)$ to vary a hundredfold. Thus

$$0.5 < k_1(\xi) < 50 \tag{6.5-4}$$

It is further assumed that although ξ can vary, the dynamics of variations encountered are slow relative to the limit cycle amplitude transient time.

It must be mentioned at this point that the simplified model for roll dynamics is chosen merely to expedite a presentation of the limit cycling adaptive-control philosophy. The general method to be discussed can be extended easily to the more complex situation of Eq. (6.5-1). This is primarily due to the fact that limit cycling control systems behave like high-gain linear systems, without the conventional problem of linear instability.

The control surfaces are driven by a nonlinear actuator, consisting of a solenoid valve with external orifice control and a hydraulic ram. The corresponding equation relating control-surface position δ to the valve input signal x can be written as

$$\dot{\delta}(\pm x) = \pm D \tag{6.5-5}$$

or in "transfer function" form,

$$\delta(s) = \frac{D}{s}\operatorname{sgn} x \tag{6.5-6}$$

The magnitude D is controlled by varying the size of the effective solenoid valve orifice. The nonlinearity considered is, in essence, an ideal relay with drive levels $\pm D$ followed by an integrator.

For the sake of subsequent reference to a specific system, the following arbitrarily chosen control-system specifications are selected:

Roll step response percent overshoot < 15 percent
Roll response time (to within ± 5 percent of final value) < 5 sec
Limit cycle amplitude at output must be constant $\leq \alpha$ deg

where α, a number related to mission requirements, need not be specified further for this example.

LIMIT CYCLING MINOR LOOP

Consider the limit cycling loop of Fig. 6.5-1, where for the moment disturbance torques are ignored. The forward path contains the hydraulic actuator and the missile airframe dynamics, and also some unspecified linear compensatory filter $H_1(s)$. The feedback element is a rate gyro with gain factor $1/k$.

With this choice of loop elements, the limit cycle frequency can be selected as desired. The equation governing the limit cycle for any loop with nonlinearity N and linear elements $HL(s)$ is

$$N_A(A)HL(j\omega_0) = -1$$

In the present case, therefore,

$$\frac{4D}{\pi A} H_1(j\omega_0) \frac{k_1(\xi)}{k} \frac{1}{(j\omega_0)^2} = -1 \tag{6.5.7}$$

which can be resolved into the real equations

$$\frac{4D}{\pi A} |H_1(j\omega_0)| \frac{k_1(\xi)}{k} \frac{1}{\omega_0^2} = 1 \tag{6.5-8}$$

and

$$\underline{/H_1(j\omega_0)} = 0 \tag{6.5-9}$$

In anticipation of a system response time of about 4 sec and in conjunction with hypothetical missile structural capabilities, $\omega_0 = 20$ radians/sec is

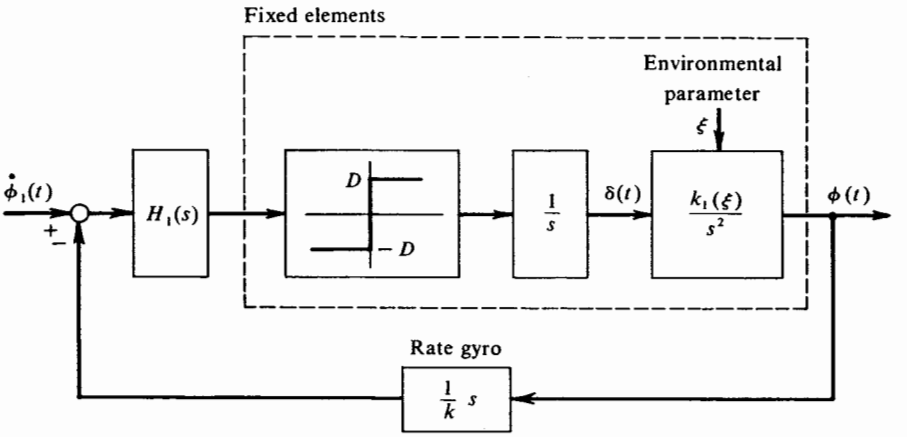


Figure 6.5-1 Limit cycling inner loop.

chosen. Clearly, to achieve a stable limit cycle, $H_1(s)$ must be a lead-lag network. It may be chosen as

$$H_1(s) = \frac{s+1}{0.1s+1}\frac{0.0025s+1}{0.025s+1} \tag{6.5-10}$$

which satisfies Eq. (6.5-9) at $\omega_0 = 20$ radians/sec. A root-locus plot for the limit cycling loop is shown in Fig. 6.5-2. The closed-loop poles marked off correspond to a value of equivalent open-loop gain equal to half that which sustains the limit cycle, in accordance with the previously established relationship for an ideal relay [Eq. (6.3-5)]. From the root-locus plot the minor-loop transfer function is determined as

$$\begin{aligned}\frac{\varphi}{\dot{\varphi}_1}(s) &\approx \frac{k}{s}\frac{(1.042)(214)(s+1)}{(s+1.042)(s^2+3.5s+214)} \\ &\approx \frac{214k}{s(s^2+3.5s+214)}\end{aligned} \tag{6.5-11}$$

The approximate cancellation of the open-loop zero by the closed-loop pole is a manifestation of the high-gain character of limit cycling control loops mentioned earlier. The pole at the origin in this transfer function is due to the feedback element. Thus the application of the theory of limit cycling systems to this limit cycling loop leads to a simple conclusion regarding its input-output dynamical properties. *Note that these properties are independent of* $k_1(\xi)$; the basis for a parameter-adaptive control system has been achieved.

ROLL POSITION OUTER LOOP

Design of the linear outer loop is a matter of selecting a series prefilter which causes satisfaction of the amplitude-ratio condition at the input to the nonlinearity over the expected class of inputs and simultaneously provides a suitable input-output transfer for the entire system. Thus $H_0(s)$ in Fig. 6.5-3 must be determined. It is a prefilter to the limit cycling loop, but internal to the roll position loop. A second-order prefilter suffices for the problem at hand. Thus

$$H_0(s) = \frac{\omega_f^2}{s^2 + 2\zeta_f\omega_f s + \omega_f^2} \tag{6.5-12}$$

To satisfy the frequency-ratio requirement we may conservatively choose $\omega_f = 0.1\omega_0 = 2$ radians/sec. That this choice is conservative follows from the fact that the prefilter output consists of forced responses and normal modes, wherein the normal modes (for this hypothetical missile problem) are assumed to have significant energy in a higher frequency band than the forced responses. Both ζ_f and k remain to be chosen. These may be

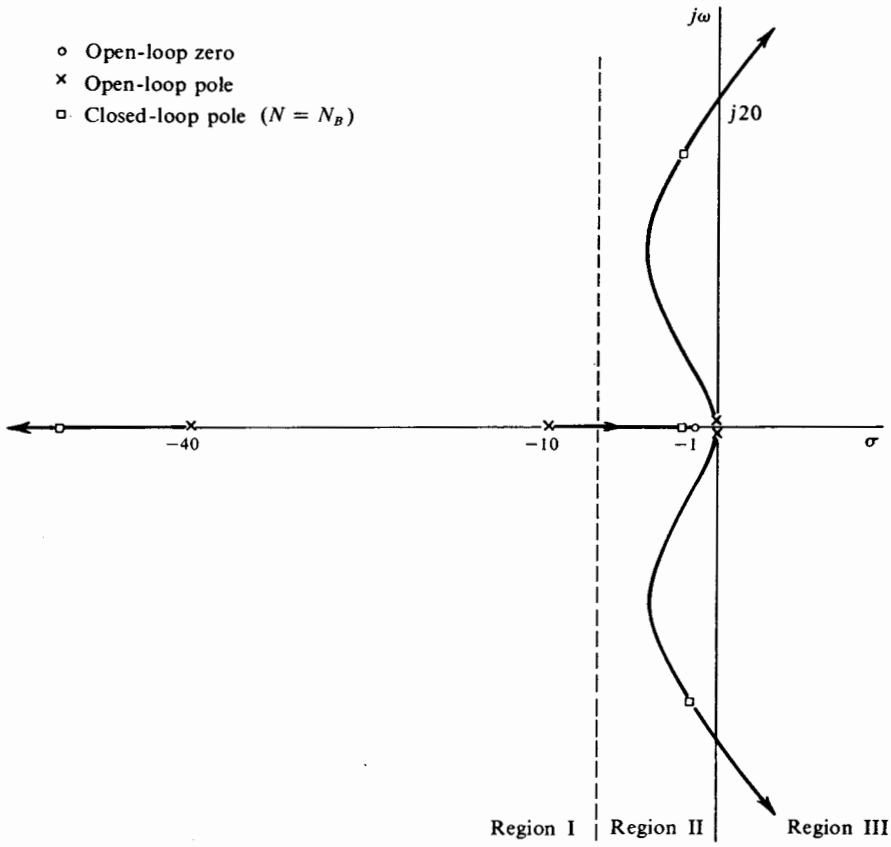


Figure 6.5-2 Minor-loop root-locus plot.

obtained by consideration of the amplitude-ratio condition. In particular, we can choose φ_{cmd} to be a step function (a reasonable choice in this roll problem), replace N by its linear equivalent, and solve for $x(t)$ in terms of the amplitude of φ_{cmd}, ζ_f, and k. It has been determined that the range of choices for ζ_f and k given by

$$\begin{aligned} \zeta_f &> 0.6 \\ 4\zeta_f &> k > 0.7 \end{aligned} \tag{6.5-13}$$

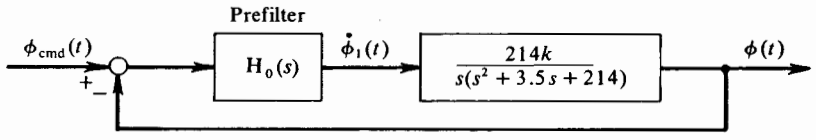


Figure 6.5-3 Roll position outer loop.

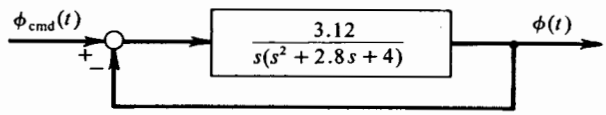


Figure 6.5-4 Equivalent overall system.

results simultaneously in system stability and satisfaction of the amplitude-ratio condition. One set of values within this range which also provides the desired system transient response is

$$\zeta_f = 0.7 \qquad k = 0.78 \tag{6.5-14}$$

The resulting equivalent overall system is shown in Fig. 6.5-4, where the complex pole pair of the inner loop has been dropped from further consideration, being well beyond the bandwidth of the system.

AMPLITUDE-REGULATING LOOP

If design is terminated at this point we shall have a system which satisfies all specifications except that of output limit cycle amplitude. In fact, as the missile sensitivity $k_1(\xi)$ varies over a 100:1 range, the limit cycle amplitude at station φ follows proportionally, whereas that at station δ (Fig. 6.5-1) remains constant. Since we have made available a variable nonlinearity (variable orifice), we are at liberty to reverse the above situation (see Fig. 6.5-5).

The active device $M[A]$ measures the output limit cycle amplitude and adjusts the valve orifice so that this limit cycle amplitude remains constant at α deg (see specifications). Details of the general requirements for construction of $M[A]$ and the resultant speed of adaptation of the overall system can be deduced from the content of Sec. 4.4. Note that $M[A]$ serves to establish limit cycle amplitude regulation at the system output. This, of course, implies that the limit cycle amplitude at the control surface varies over a 100:1 range. Therefore D is chosen so that, at the smallest value for $k_1(\xi)$, the limit cycle amplitude at δ is acceptable. As $k_1(\xi)$ increases, this limit cycle amplitude decreases, thus always remaining acceptable. Setting D also relates to the amplitude-ratio condition, with a slight circular nature to this problem arising. A solution is reached when the amplitude-ratio condition at x is not violated for the largest expected system transient input, when the limit cycle at station δ is acceptable, and when the overall system response fits specifications. Thereafter the system response is invariant.

Note that the regulation of the limit cycle amplitude in no way alters the

system input-output behavior as long as the amplitude-ratio condition is always satisfied.

We may digress a moment to discuss the disturbance response properties of the control system at this point. Equation (6.5-3) indicates the manner in which disturbance torques enter the system. Figure 6.5-5 details this situation. To determine the effect of disturbance torques on roll angle it is necessary to specify, first, the energy distribution in $L_D(s)$. This is due to the nonlinear filtering qualities of an apparently "dithered" relay seen by any signal. For low-frequency disturbance inputs which also do not violate the amplitude-ratio condition at x, the system behaves as a linear feedback system, and its transfer characteristic may be so determined. The steady-state error, for example, in response to any constant disturbance is zero. For high-frequency disturbance inputs the limit cycling system appears open-loop, so that

$$\frac{\varphi}{L_D}(s) \approx \frac{k_2(\xi)}{s^2} \tag{6.5-15}$$

For exceedingly large transient disturbances the limit cycle may temporarily terminate with the "relay" hard over (i.e., control surface slewing at maximum rate in one direction to overcome the disturbance).

EXPERIMENTAL RESULTS

An analog computer was used in a study of the system of Fig. 6.5-5. The function of $M[A]$ was performed manually, according to the same rule which would govern the actual circuit. Figure 6.5-6 shows the system roll transient response for a number of conditions over the 100:1 range in $k_1(\xi)$. Note that in all cases the time to peak is approximately 2.8 sec and the overshoot is 12 percent. For the system of Fig. 6.5-4 a purely linear calculation predicts corresponding values of 2.9 sec and 12.5 percent. Thus the argument that the overall system behaves linearly is well supported. Further, observe that the limit cycle amplitude at the output is so small as to be barely discernible on the analog-computer recordings. Figure 6.5-7 shows δ at intervals over the same range. Note the ordinate scales. It is seen that the correlation between theory and experiment is excellent.

Finally, recall that the time-domain specifications were chosen arbitrarily, merely to allow numerical reference throughout the control-system design. The ultimate capability of the hypothetical airframe is by no means exhausted. In the final analysis this depends upon structural considerations and control-surface slewing capability. In fact, recognizing that the control-surface

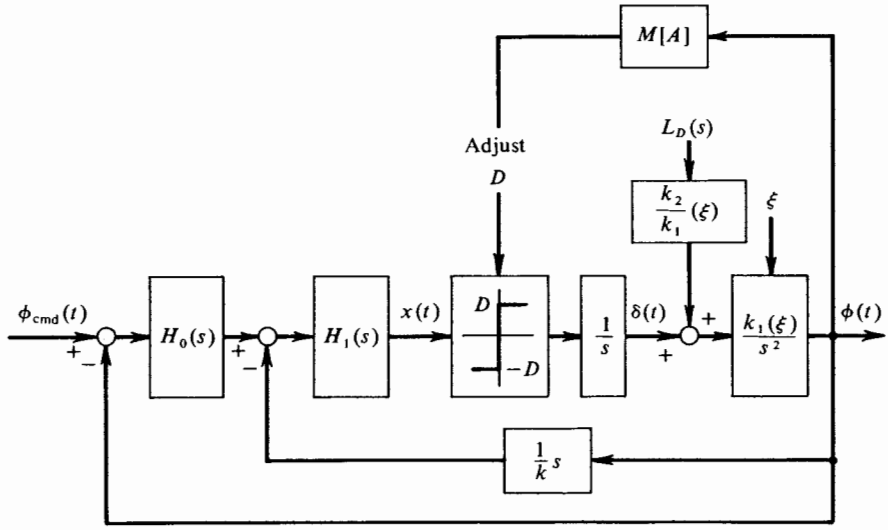


Figure 6.5-5 Proposed control configuration.

limit cycle amplitude is largest when $k_1(\xi)$ is smallest, and that this condition may exist for only a small part of the flight, quite significant decreases in system response time may be obtained. In any case, the simplicity in design of this system is clear, indicating the potential practicality of its application to the missile roll-stabilization problem.

6.6 LIMIT CYCLES IN SYSTEMS WITH AN ASYMMETRIC NONLINEARITY

A very useful property of the DIDF formulation is the means by which systems with an asymmetric[1] nonlinearity can be studied. As we shall presently see, this application is quite natural, and conceivably could have accounted for the DIDF development in the first place.

Consider the feedback system of Fig. 6.6-1 in which N is an asymmetric nonlinear element and $L(s)$ is the loop linear part. If this loop is in a limit cycle state, it is clear that $x(t)$ will contain a bias term [for our present purposes we choose $r(t) = 0$]. This follows since, in the assumed absence of any bias in $x(t)$, $y(t)$ will indeed develop a bias and contradict this assumption. Such behavior, of course, results solely from our assertion that the nonlinearity is asymmetric. In a symmetric nonlinearity system no biases appear, a fact which is implicit in the formulation of the limit cycle theories of previous chapters.

[1] By asymmetric is meant one which is not odd; that is, $y(x) \neq -y(-x)$.

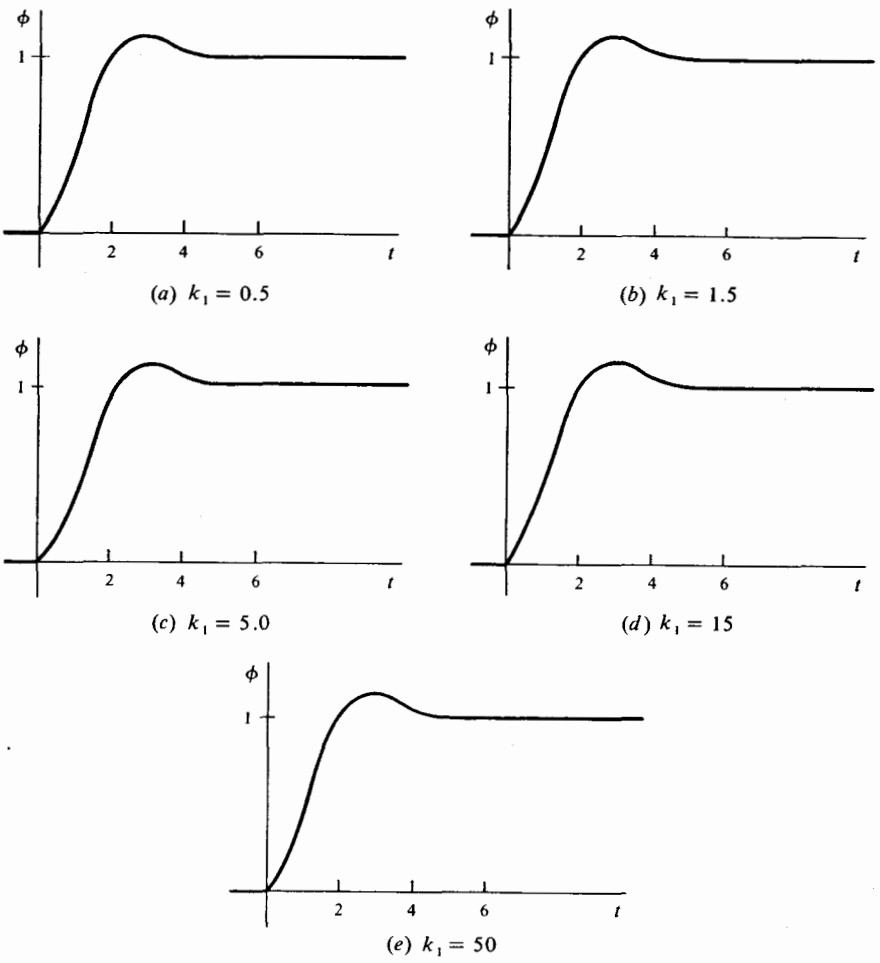


Figure 6.5-6 *Missile roll responses taken over the entire flight envelope.*

For a sinusoid at $x(t)$ to propagate unattenuated around the loop requires

$$N_A(A,B,\omega_0)L(j\omega_0) = -1 \tag{6.6-1}$$

as in past limit cycle formulations. In addition, however, it is required that a bias (B) at $x(t)$ also propagate identically around the loop. Thus

$$BN_B(A,B,\omega_0)L(j0) = -B \tag{6.6-2}$$

This identity is automatically satisfied if $B = 0$, an uncommon situation, and hence is of little interest in that case. For $B \neq 0$, however, we must have

$$N_B(A,B,\omega_0)L(j0) = -1 \tag{6.6-3}$$

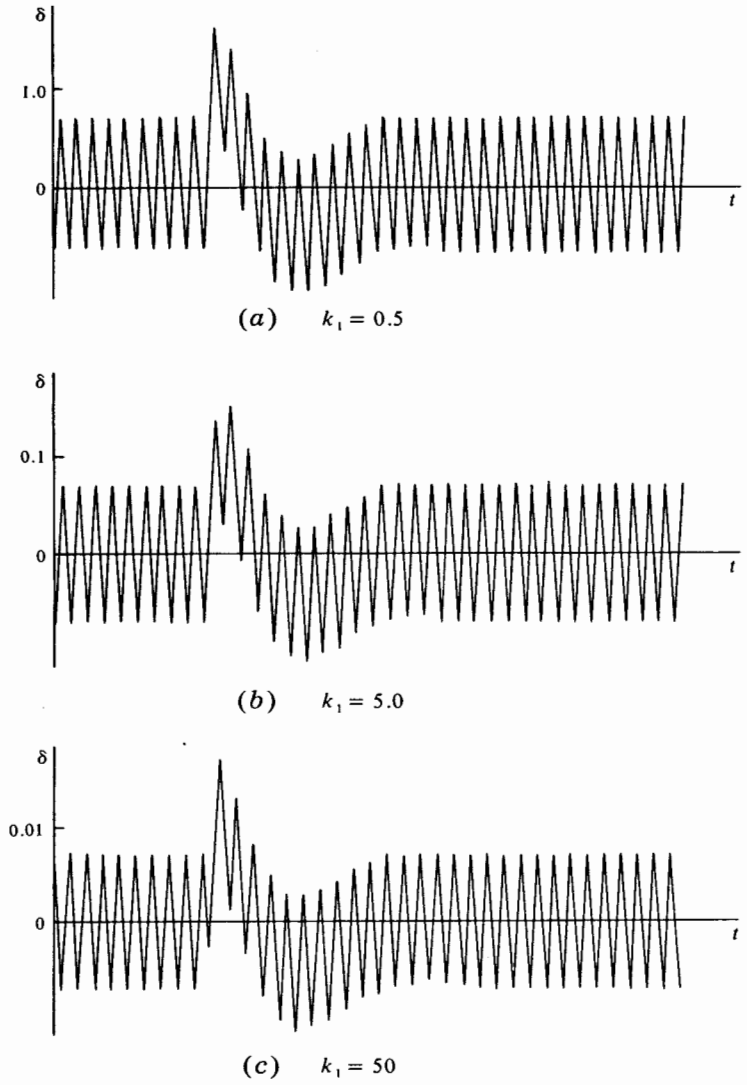


Figure 6.5-7 *Control-surface responses taken over the entire flight envelope.*

the second of *two* conditions which must be fulfilled in order that the system of Fig. 6.6-1 sustain a limit cycle. $L(j0)$ denotes, of course, the static sensitivity (dc gain) of the linear elements.

It is true, in general, that the periodic output of an asymmetric nonlinearity in a limit cycling system contains *even* harmonics. In particular, some *second* harmonic may be present. To whatever extent this is true, the filter

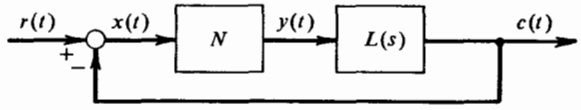


Figure 6.6-1 System with an asymmetric nonlinearity.

hypothesis need obviously be altered in the direction of requiring *more* filtering of the nonlinearity output for continued validity of the DIDF linearization. Otherwise, all remains as before.

When $L(s)$ has no open-loop integrations (poles at the origin), $L(j0)$ is finite and the application of Eq. (6.6-3) is straightforward. If $L(s)$ has one or more open-loop integrations, $|L(j0)| \to \infty$, and we require $N_B(A,B,\omega_0) = 0$ [alternatively, $BN_B(A,B,\omega_0) = 0$ since $B \neq 0$] as the only possible solution of Eq. (6.6-3). If, on the other hand, $L(s)$ has one or more open-loop differentiations (zeros at the origin), the only possible means for satisfying the limit cycle conditions are indicated by Eq. (6.6-2), which calls for $B = 0$. In each of these cases, stability of the indicated bias value must be determined. Table 6.6-1 summarizes the possible circumstances.

TABLE 6.6-1 REQUIREMENTS ON B FOR A LIMIT CYCLE TO EXIST†

$L(s)$		Requirement
$s^n L_1(s)$	$n \geq 1$	$B = 0$
$L_1(s)$		$N_B(A,B,\omega_0)L_1(j0) = -1$
$\dfrac{L_1(s)}{s^m}$	$m \geq 1$	$BN_B(A,B,\omega_0) = 0$

† $L_1(s)$ has neither poles nor zeros at the origin.

A simple example serves to clarify the use of this table.

Example 6.6-1 Determine the limit cycle state of the system in Fig. 6.6-2.

The asymmetric nonlinearity is effectively a biased-output ideal relay, for which $N_A(A,B)$ and $N_B(A,B)$ appear in Appendix C. For $|B/A| \leq 1$

$$N_A(A,B) = \frac{2D}{\pi A}\sqrt{1 - \left(\frac{B}{A}\right)^2}$$

$$N_B(A,B) = \frac{D}{2\pi B}\left(\pi + 2\sin^{-1}\frac{B}{A}\right)$$

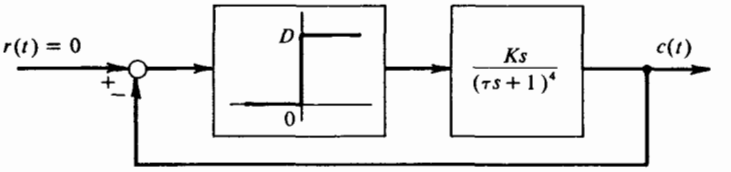


Figure 6.6-2 Example system with a biased-output ideal-relay nonlinearity.

Table 6.6-1 evidences that $B = 0$ is a limit cycle requirement. Hence, from Eq. (6.6-1) it follows that

$$\frac{2D}{\pi A}\frac{Kj\omega_0}{(j\tau\omega_0 + 1)^4} = -1$$

Solution yields

$$\omega_0 = \frac{2.41}{\tau}$$

and

$$A = 0.033\frac{DK}{\tau}$$

It remains to be shown that the condition of zero nonlinearity input bias is a stable one. To do this, assume a perturbation about zero of the input bias, and the associated perturbation about $D/2$ of the output bias. This gives the incremental-input describing function applicable to the present example.

$$N_B(A,B=0) \triangleq \lim_{\Delta B\to 0}\frac{\Delta B N_B(A,\Delta B) - D/2}{\Delta B}$$

$$= \frac{D}{\pi A}\lim_{\Delta B\to 0}\frac{\sin^{-1}(\Delta B/A)}{\Delta B/A}$$

$$= \frac{D}{\pi A}$$

$$= \frac{9.65\tau}{K}$$

We are thus interested in stability of the system illustrated in Fig. 6.6-3. The open-loop elements of this *linear* feedback system are

$$\text{OLTF}(s) = 9.65\frac{\tau s}{(\tau s + 1)^4}$$

By use of linear analysis it is readily verified that this system is indeed stable. Hence, and in summary, the limit cycle state determined above is stable. An analog simulation of this system bears out these findings, and indicates solution accuracies of about 5 percent.

If the linear elements of this example are replaced by either

$$\frac{K}{(\tau s + 1)^3} \quad \text{or} \quad \frac{K}{s(\tau s + 1)^2}$$

the DIDF analysis leads to incompatible requirements, from which it is to be concluded that a limit cycle is not possible. Again, analog simulation substantiates these conclusions.

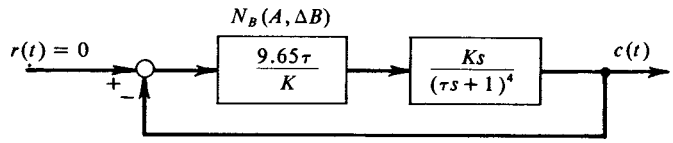


Figure 6.6-3 Bias stability study of example system.

6.7 ARTIFICIAL DITHER AND SIGNAL STABILIZATION

The vehicle by which DIDF formulation, calculation, and usage have been conveyed thus far in this chapter has been the limit cycling system. Other interpretations of the bias and sinusoidal part of the nonlinearity input signal exist, and they account for quite different behavioral aspects of nonlinear systems. Consider, for example, a non-limit-cycling system with an input

$$r(t) = r_B(t) + R \sin \omega_r t$$

where $|\dot{r}_B(t)| \ll R\omega_r/2\pi$. Provided that the application of this signal does not stimulate a systematic oscillation and that the amplitude of the low-frequency part of the nonlinearity input is less than that of the sinusoidal part, DIDF linearization of the system nonlinearity provides a means for determining the forced response behavior of the system. Although it is an easy matter to check the amplitude-ratio requirement, it is quite difficult to determine whether the input *causes* the system to oscillate at frequencies other than ω_r. Describing function verification would call for a nonlinearity linearization for inputs comprised of a bias and *two* sinusoidal inputs. In theory, this can certainly be achieved. Practically, this may go beyond the point of diminishing returns. For this reason attention is confined to situations where simple DIDF linearization is useful by itself.

ARTIFICIAL DITHER

The employment of high-frequency signal injection for the purpose of altering the *apparent* characteristics of a nonlinearity in a closed-loop system is an exceedingly useful compensatory device. The high-frequency signal is referred to as *artificial dither*, or just *dither*.[1] The technique of dithering for linearizing purposes has been known for some time.

Artificial dither implies intentionally applied dither. Unintentional introduction of dither into control systems, however, is not at all uncommon. Examples are 60- or 400-cps electrical pickup in an instrument servo, and

[1] In this chapter all dither waveforms considered are periodic. The use of random dither is discussed in Chap. 7.

mechanical vibration in a missile control-surface servo. Under the circumstances to be delineated presently, such unintentional dither can be readily accounted for as well.

In what follows it is assumed that the loop linear elements attenuate the high-frequency dither to the point where only an insignificant dither frequency signal makes the return trip to its originating place. Three types of dither are considered separately.

Triangular-wave dither Consider the application of symmetric triangular-wave dither to an *arbitrary* memoryless odd saturating nonlinearity (Fig. 6.7-1*a*). Let $e(t)$ represent the total nonlinearity input, comprised of signal $x(t)$ and dither $d(t)$. This symmetric nonlinearity is defined by

$$\begin{aligned} e &\leq -\delta & y &= -D \\ -\delta < e &< \delta & y(e) &= -y(-e), \text{ otherwise arbitrary} \\ e &\geq \delta & y &= D \end{aligned} \tag{6.7-1}$$

Input and output waveforms are shown in Fig. 6.7-2, where for input values in the range $-\delta < e < \delta$ the (arbitrary) output is shown dotted. Since the

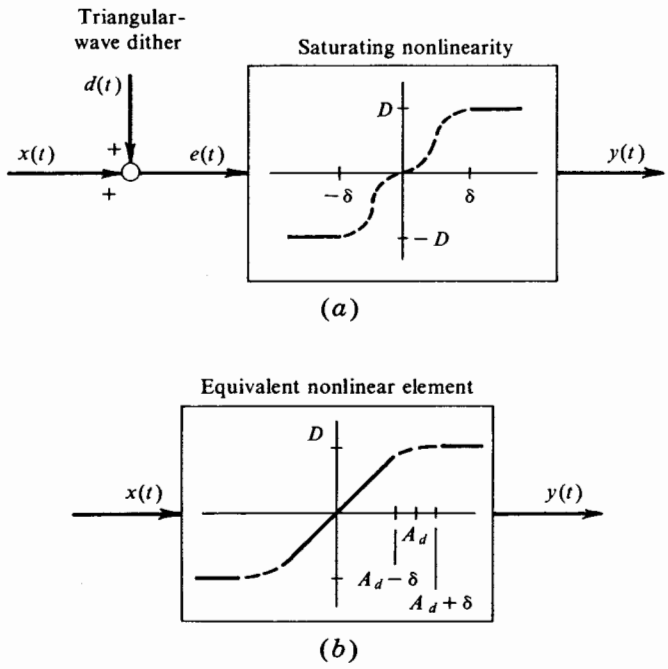


Figure 6.7-1 *(a) Application of triangular-wave dither to an arbitrary odd saturating nonlinearity and (b) the equivalent nonlinear element.*

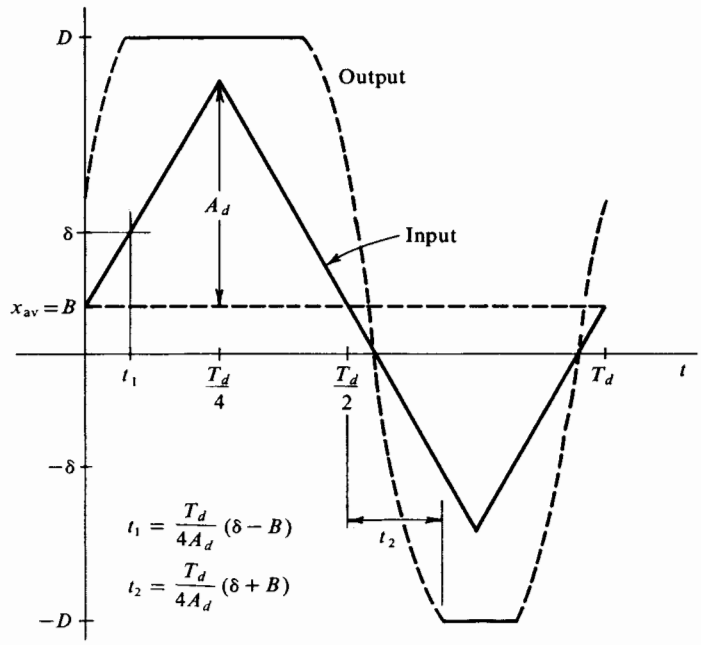


Figure 6.7-2 *Input and output waveforms for an odd saturating nonlinearity with triangular-wave dither.*

dither frequency is by definition significantly higher than the frequencies of interest in x, the input signal is modeled as a bias over one dither cycle.

The signal gain of this device is, by analogy with previous DIDF formulations,

$$N_B(A_d,B) = \frac{y_{av}}{B} \tag{6.7-2}$$

In terms of the parameters of the illustration, the average output is given by (T_d = dither period)

$$y_{av} = \frac{1}{T_d}\left[D\left(\frac{T_d}{2} - 2t_1\right) - D\left(\frac{T_d}{2} - 2t_2\right)\right]$$

$$= \frac{BD}{A_d} \tag{6.7-3}$$

Notice that no account need be taken of the dotted portion of the nonlinearity output, since over a complete cycle it has an average value of zero. It is for this reason that arbitrary nonlinearity characteristics can be admitted,

provided they are odd. Hence, for $|B| < A_d - \delta$, the signal gain of this device is

$$N_B(A_d,B) = \frac{D}{A_d} \tag{6.7-4}$$

which, remarkably, is independent of δ. The inequality requirement on B points out that we have assumed the nonlinearity output to saturate positively and negatively during every dither cycle. For values of B in the range

$$A_d - \delta < |B| < A_d + \delta$$

the signal gain depends upon the particular nonlinearity under consideration. Values of B satisfying

$$|B| > A_d + \delta$$

lead to a constantly saturated nonlinearity output. Figure 6.7-1*b* summarizes these results by illustrating the resulting equivalent nonlinear element. Of great consequence is the fact that *the equivalent nonlinear element appears absolutely linear about the origin, irrespective of the shape of the original nonlinearity about the origin.* The range of linearity can be increased simply by increasing A_d, accompanied, of course, by a decrease in linear gain. For sufficiently large input signals the equivalent nonlinear element saturates at the output saturation value of the original nonlinearity.

By very similar reasoning it can be demonstrated that these results carry over to the entire class of odd saturating nonlinearities with memory.

Square-wave dither Provided that the dither plus bias drive the output to saturation on the positive *and* negative parts of every dither cycle, the average output of the nonlinear element is identically zero for all nonlinearities given by Eqs. (6.7-1). Continuing as before yields, for the arbitrary symmetric nonlinearity with square-wave dither shown in Fig. 6.7-3*a*, the equivalent nonlinearity of Fig. 6.7-3*b*. A_d is the square-wave amplitude. As before, dotted line segments refer to portions of the characteristics which vary from one nonlinearity to another.

In this case, the noteworthy fact is that *the equivalent nonlinear element possesses a dead zone about the origin, regardless of the shape of the original nonlinearity about the origin.* Again, the results apply to nonlinearities with memory, as well as to those without memory.

Sinusoidal dither The case of sinusoidal dither is identical with that for which the DIDF calculations of previous sections were performed. No generalization can be made here, except that, for sufficiently large A_d, the effect of sinusoidal dither is similar to that of triangular-wave dither in that *all* nonlinearities are linearized about the origin. In the case of sinusoidal dither, the assumption of negligible fed-back dither signal need not be held.

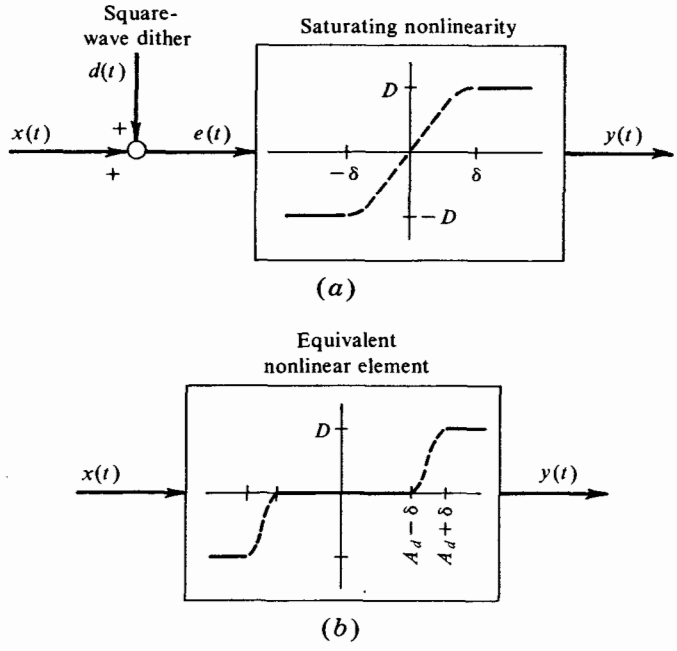


Figure 6.7-3 (*a*) *Application of square-wave dither to an arbitrary odd saturating nonlinearity and* (*b*) *the equivalent nonlinear element.*

With additional labor, but with no new analytic means required, the feedback sinusoid can be included in determining the total sinusoidal input to the nonlinearity.

Additional remarks It follows from the discussion of the effect of artificial dither that the ideal-relay characteristic, for example, can be manipulated as shown in Fig. 6.7-4. Similar results hold for many other nonlinearities. Thus a practical means for altering the properties of a control loop is that of simply changing the *waveform* of an additive dither signal. In two of the three cases shown, the effect of this dither is to *linearize* the nonlinearity for small input signals, the triangular dither providing a larger linear region than the sinusoidal dither. This technique is profitably employed in the compensation of nonlinear control systems, often for the purpose of linearizing the highly nonlinear force-velocity effects due to friction. Since the apparent nonlinearity characteristics are dependent upon A_d, the use of dither of fixed waveshape but variable amplitude allows for continuous *abscissa scaling* of each of the apparent nonlinearity characteristics illustrated. In a following servo, for example, if A_d were instrumented to be a function nearly inversely proportional to the servo error signal, the result would be a high loop gain

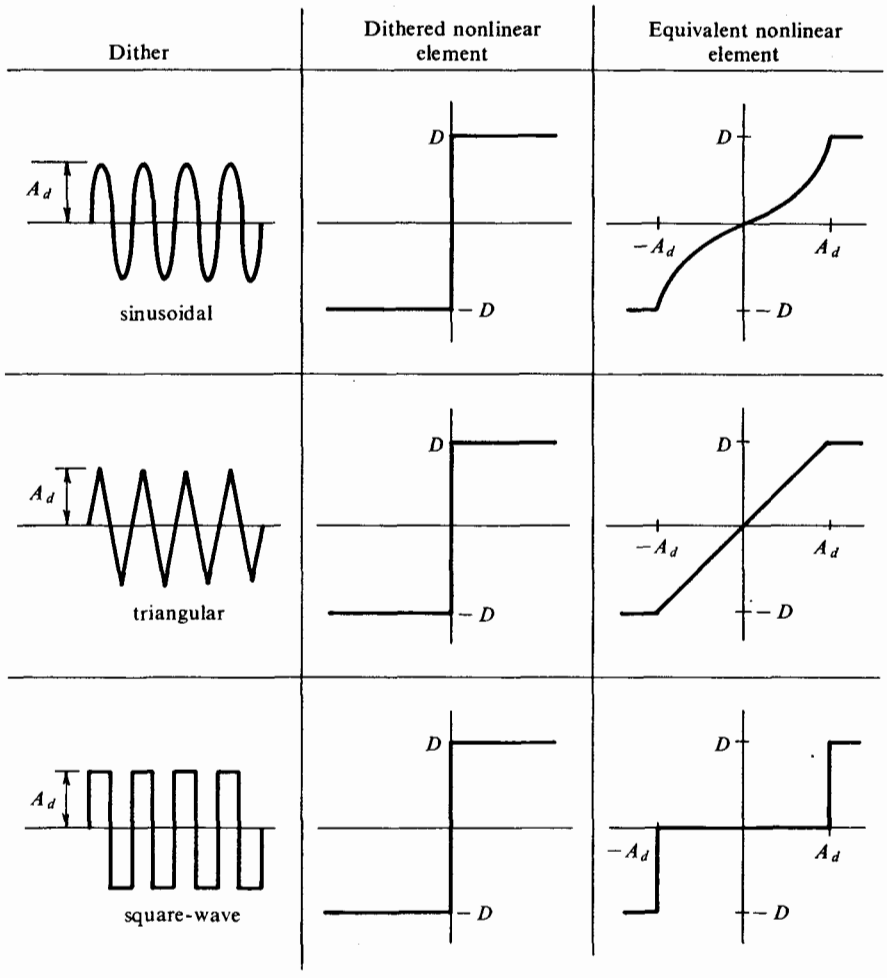


Figure 6.7-4 Ideal relay with three types of dither and the associated equivalent nonlinear characteristics.

at large error signals and a lower loop gain at small error signals. In such a way one can achieve the continuous analog of certain dual-mode servos which attempt to give near-maximum-effort performance without displaying a limit cycle.

SOME APPLICATIONS FOR ARTIFICIAL DITHER

In addition to the more common uses of artificial dither such as linearization of coulomb friction, threshold, and hysteresis effects, Naslin (Ref. 15) describes several other interesting applications.

Consider the use of dither for the purpose of providing a proportional torque zone in a relay-controlled split-field series motor (see Fig. 6.7-5). Assuming the motor to be fed by a constant current source (thus minimizing back emf effects), its output torque can take on only two values, $\pm T$, depending upon the state of the switch. The switch has two additive inputs, a signal $x_B(t)$ and dither $d(t)$. By previous DIDF arguments the average output torque is seen to be a linear function of the input signal $x_B(t)$ for small values of $|x_B/A_d|$. This concept of torque control can also be applied to other forms of energy. Basically, it amounts to a form of pulse-width modulation. Observe that grid-controlled rectifiers and some magnetic amplifiers operate in the very same manner.

Next, consider the use of dither for controlling small guided missiles by means of spoilers. Spoilers are vanes that disturb normal airfoil airflow, thus producing aerodynamic forces. The average aerodynamic force becomes a linear function of the input signal as the result of relay-controlling spoiler position such that it is either fully retracted or fully protruding, and supplying a linearizing dither signal to the relay, in addition to a signal input. Hence a linear region of control is effected.

Another application involves a position servo incorporating a photocell error detector with an *absolute-value* characteristic. This device was built for the purpose of providing an instrumented signal representation of the height of a mercury column, in a form suitable for telemetering. The servomanometer functional diagram is shown in Fig. 6.7-6. A lead-screw positioned carriage houses a photocell which detects whatever light flux (shown dotted) passes through the mercury-column-supported slit and over or under the 50-cps sinusoidally oscillating carriage-mounted diaphragm. The photocell output is filtered for recovery of whatever 50-cps content it contains, then amplified and fed to the control winding of a two-phase motor

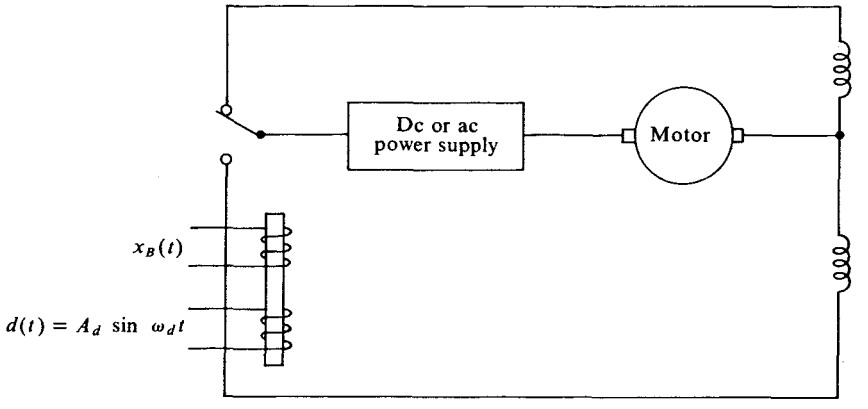


Figure 6.7-5 *Relay-controlled split-field series motor.* (*Naslin, Ref. 15.*)

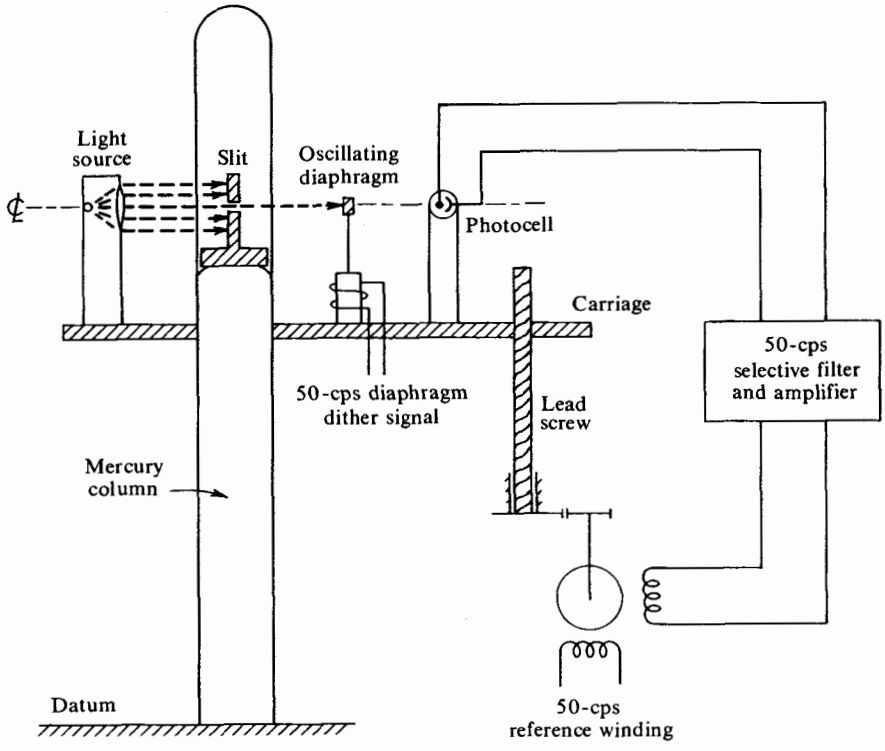


Figure 6.7-6 Servomanometer. (Naslin, Ref. 15.)

which drives the lead screw. Lead-screw angular position is then the instrumented quantity which is sensed by a shaft-angle encoder and telemetered.

To understand the behavior of this system it is necessary to determine the relationship between photocell output and slit-position input (see Fig. 6.7-7). Specializing DIDF calculations already presented for the two-segment piecewise-linear asymmetric nonlinearity to the case of an absolute-value characteristic ($m_1 = -m_2 = m$), and assuming $|e/A_d| < 1$, where $e = r - c$, yields [Eq. (6.2-11)]

$$N_A(A_d,e) = \frac{2m}{\pi}\left[\sin^{-1}\left(\frac{e}{A_d}\right) + \frac{e}{A_d}\sqrt{1 - \left(\frac{e}{A_d}\right)^2}\right]$$

$$\approx \frac{4me}{\pi A_d} \tag{6.7-5}$$

and [Eq. (6.2-14)]

$$N_B(A_d,e) = \frac{2m}{\pi}\left[\sin^{-1}\left(\frac{e}{A_d}\right) + \frac{A_d}{e}\sqrt{1-\left(\frac{e}{A_d}\right)^2}\right]$$
$$\approx \frac{2mA_d}{\pi e} \tag{6.7-6}$$

Hence the photocell output is given by

$$y(t) \approx \frac{2mA_d}{\pi} + \frac{4me}{\pi}\sin 100\pi t \tag{6.7-7}$$

where second-order and double- and higher-frequency terms have been dropped. Equation (6.7-7) exhibits a most interesting effect, namely, that *the amplitude of the first-harmonic term in $y(t)$ is directly proportional to the positioning error e.* The photocell plus dither act, in effect, as a modulator. Operation of the remainder of the system is clear, and need not be discussed further. The overall system accuracy is reported to be 0.1 mm of vertical mercury motion, or 0.002 psi.

Other diverse applications for artificial dither include extremal control, adaptive control, and signal stabilization. The last-mentioned application is briefly discussed next.

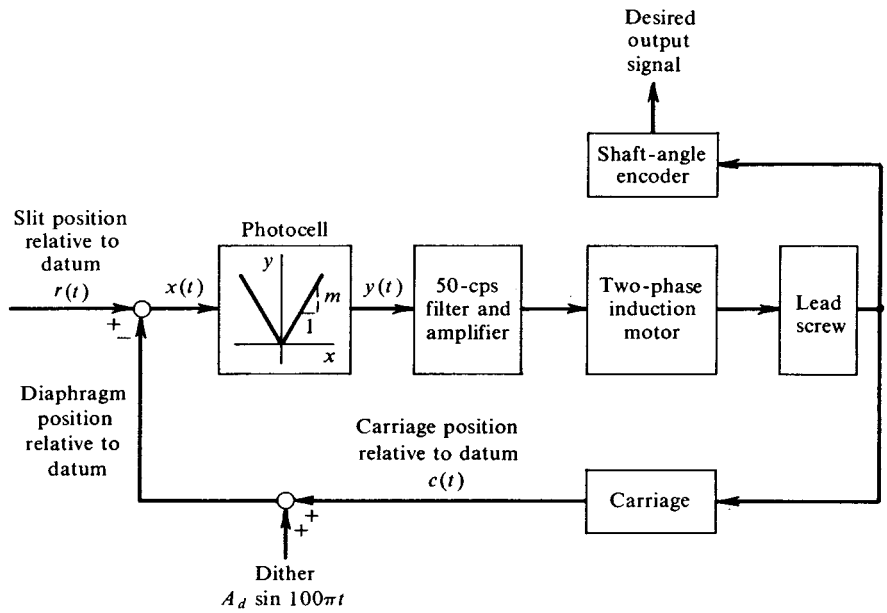


Figure 6.7-7 Block diagram of the servomanometer.

SIGNAL STABILIZATION

Since the injection of dither has the *effect* of altering nonlinear characteristics in closed-loop systems, it is reasonable to expect that limit cycles in nonlinear systems can be turned on, altered, turned off, and, in general, controlled by proper choice of dither waveshape. The use of dither to turn limit cycles off is referred to as *signal stabilization.* It has been extensively investigated by Oldenburger and his students (Refs. 16–19), among the first to discover this phenomenon experimentally and, subsequently, to provide analytical justification. They point out that the analytical justification for signal stabilization differs from that given by Minorsky (Ref. 13) relating to the asynchronous quenching of systems characterized by nonlinear differential equations.

Investigation of signal stabilization via describing function theory can be executed as a two-stage process. First, by use of DIDF theory, the dither and original nonlinear element are replaced with an equivalent nonlinear element, whose form implicitly accounts for the presence of dither, but which no longer explicitly displays the dither signal. Second, the resulting system is made the object of a DF analysis to reveal the presence or absence of a limit cycle. This rather simple investigation scheme is demonstrated by the following example.

Example 6.7-1 Find the minimum sinusoidal dither amplitude $A_{d,\min}$ for which the system of Fig. 6.7-8*a* just fails to develop a limit cycle, given that $\omega_d \gg 1$.[1]

First, the relay with dither is replaced by its equivalent nonlinear element (signal DIDF) (Fig. 6.7-8*b*). From this point analysis proceeds by determination of the DF for this equivalent nonlinear element and subsequent performance of a DF limit cycle analysis of the resulting system. Note that the limit cycle, if any, plays the role previously played by system error *signals*, and the dither plays the role otherwise belonging to a limit cycle.

The DF of the signal DIDF is determined as follows ($k = A/A_d < 1$):

$$N(A) = \frac{4}{\pi A} \int_0^{\pi/2} \frac{2D}{\pi} \sin^{-1}(k \sin \psi) \sin \psi \, d\psi$$

Integrating by parts and grouping terms yields

$$N(A) = \frac{8D}{\pi^2 A} \left(\frac{1}{k} \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2 \psi} \, d\psi + \frac{k^2 - 1}{k} \int_0^{\pi/2} \frac{d\psi}{\sqrt{1 - k^2 \sin^2 \psi}} \right)$$

$$= \frac{8D}{\pi^2 A_d k^2} [E(k) + (k^2 - 1)K(k)] \tag{6.7-8}$$

where $E(k)$ and $K(k)$ denote the complete elliptic integrals of the first and second kind, respectively [Eq. (5.1-18)].

[1] This system was studied by Oldenburger and Boyer (Ref. 19), although they did not determine the required DF analytically. Their results are, of course, identical with those given here.

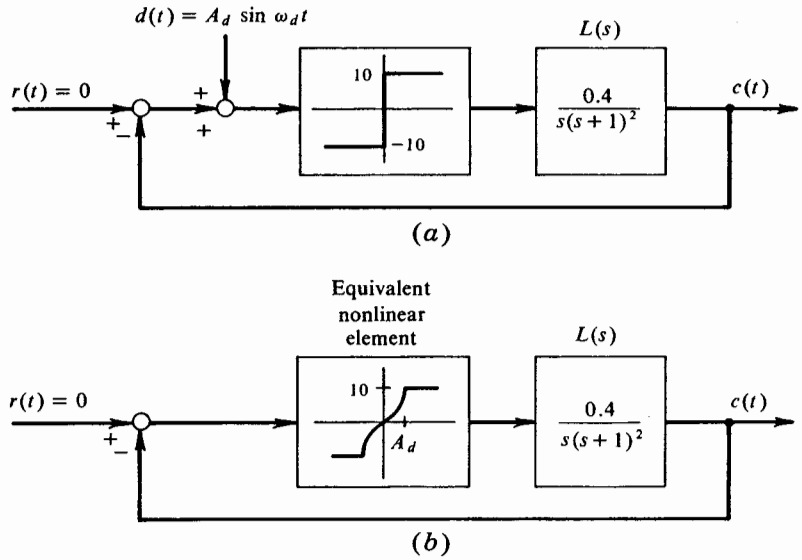


Figure 6.7-8 *(a) Relay control system with stabilizing signal added. (b) Equivalent nonlinear system.*

Similarly, it is readily shown that for $k > 1$ the DF is given by

$$N(A) = \frac{8D}{\pi^2 A_d k} E\left(\frac{1}{k}\right) \tag{6.7-9}$$

A normalized plot of $N(A)$ over a range of k is shown in Fig. 6.7-9. Since all devices which ultimately saturate at fixed levels possess DFs which are asymptotic to the ideal-relay DF for large inputs, this curve is included for reference.

Returning to the problem, it is to be observed that any limit cycle which takes place must do so at a frequency $\omega_0 = 1$,[1] where the plant phase shift is $-180°$. At this frequency the remaining condition for a limit cycle is

$$N(A)\,|L(j1)| = 1$$

Since $|L(j1)| = 0.2$, it follows that

$$N(A) = 5$$

Considering $-1/N(A)$ as plotted in the amplitude-phase plane, it is clear that the lowest point of this locus corresponds to the peak in the curve of Fig. 6.7-9. So adjusting this peak such that the $-1/N(A)$ locus never cuts the $L(j\omega)$ locus guarantees the absence of a limit cycle (to whatever extent one is willing to make quantitative guarantees based on approximate analysis). The required condition is

$$N_{\max} < 5 \qquad \text{or} \qquad \left(\frac{A_d N}{D}\right)_{\max} < \frac{5A_d}{D}$$

[1] For which reason the dither frequency was specified as greater than 1; that is, $\omega_d \gg \omega_0$. So one can neglect the fed-back dither when computing the equivalent nonlinearity.

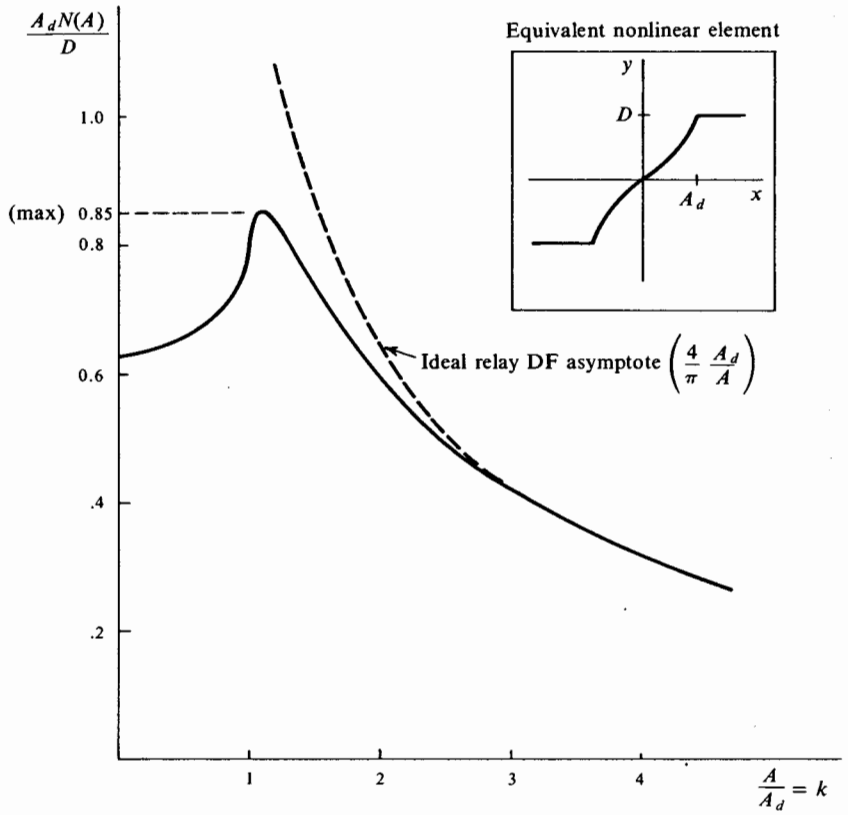


Figure 6.7-9 Equivalent nonlinear element DF for a sinusoidally dithered ideal-relay characteristic.

From the DF curve, we find $(A_d N/D)_{\max} = 0.85$; and from the problem statement $D = 10$; whence the condition on A_d for signal stabilization is

$$A_d > \frac{10(0.85)}{5} = 1.7 \text{ units}$$

or

$$A_{d,\min} = 1.7 \text{ units} \tag{6.7-10}$$

For values of A_d between 1.26 and 1.7 units, two limit cycles are indicated, of which only that of larger amplitude is stable. For values of A_d less than 1.26 units, a single stable limit cycle occurs.

Analysis as presented in the foregoing example is subject to the usual DF and DIDF limitations, such as the assumption of a single time-invariant loop nonlinearity, the absence of nonlinearity subharmonic generation, and satisfaction of the filter hypothesis. In addition, it is a convenience if only a negligible amount of dither returns via the feedback loop to the nonlinearity

input. As a rule of thumb, Oldenburger et al. suggest that the dither frequency be at least 10 times the highest possible limit cycle frequency, an assumption that is readily verified during analysis. This is roughly the same rule of thumb which ought to be used in design of a limit cycling control system, where the frequency ratio of 10 there refers to limit cycle frequency over highest significant input frequency.

The analytic study of signal stabilization as described above is contingent upon our ability to determine DFs for the equivalent nonlinear element under consideration. No difficulty is likely to arise in obtaining the equivalent nonlinear element itself. The DF can then be calculated by using the approximation techniques of Sec. 2.6. Occasionally, the DF can be readily determined analytically. This is shown for the ideal relay in Example 6.7-1, and can also be done for odd polynomial nonlinearities. In the case of a sinusoidally dithered cubic characteristic, for instance, the equivalent nonlinear element is given by [Eq. (6.2-27)]

$$y = \tfrac{3}{2}A_d^2 x + x^3 \tag{6.7-11}$$

Employing Eq. (2.3-21) yields immediately the DF as

$$N(A) = \tfrac{3}{2}A_d^2 + \tfrac{3}{4}A^2 \tag{6.7-12}$$

6.8 TSIDF CALCULATION VIA THE DF OF A DIDF

The reader may have already noticed that the DF calculations made in the previous section yielded the exact TSIDF linearization for the ideal-relay and cubic characteristics. This is by no means just coincidence; one *can* compute the TSIDF for a nonlinearity by obtaining the DF of the equivalent nonlinear element. This is readily proved.

TSIDF CALCULATION

We have seen the output of a nonlinearity expanded in a double Fourier series (Sec. 5.1). In the case of single-valued frequency-independent nonlinearities with non-harmonically-related input sinusoids, the TSIDF for one of the two input sinusoids is given by [Eq. (5.1-15)]

$$\begin{aligned} N_B(A,B) &= \frac{1}{2\pi^2 B}\int_{-\pi}^{\pi}\int y(A\sin\psi_1 + B\sin\psi_2)\sin\psi_2\, d\psi_1\, d\psi_2 \\ &= \frac{1}{\pi B}\int_{-\pi}^{\pi}\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} y(A\sin\psi_1 + B\sin\psi_2)\, d\psi_1\right]\sin\psi_2\, d\psi_2 \end{aligned} \tag{6.8-1}$$

Notice that ψ_2 is held *constant* in the integral within brackets; the argument of y is equivalently a sinusoid plus a bias. In DIDF notation, this bracketed term is represented as $BN_B(A,B)$, the bias portion of the nonlinearity output. Alternatively, it is the output of the equivalent nonlinear element. The remainder of Eq. (6.8-1) is, of course, the DF formulation. Hence the TSIDF is indeed the DF of the equivalent nonlinear element! This proof must be restricted to the case of non-harmonically-related input sinusoids, although it can be generalized to include frequency-dependent nonlinearities. This is, in fact, a special case of the general property of *any other* independent inputs expressed in Eq. (1.5-41).

TSIDF APPROXIMATION

It is of some interest to determine the circumstances under which the DIDF itself is a suitable approximation to the TSIDF. The conditions under which this use can be made of the DIDF surely cannot be stated very generally. Fortunately, it is possible to determine qualitatively the range of usefulness of this approximation without doing the more tedious calculation of the TSIDF.

Two approaches to this problem appear in Ref. 1. We note the result here, without laboring through the related mathematics. It is, simply, that if the DIDFs $N_A(A,B)$ and $N_B(A,B)$ are nearly independent of B for some range of $|B| \leq A$, then the approximate TSIDFs given by $N_A(A,0)$ and $N_B(A,0)$ are valid in this range. Since these are seen to be the leading terms in the TSIDF power-series expansions of Sec. 5.1, this approximation warrants no further discussion.

6.9 BASIS FOR HIGHER-ORDER APPROXIMATIONS

One can obtain higher-order DIDF approximations in several ways. For nonlinearities which are nonanalytic, a procedure entirely equivalent to the refined DF approximation of Sec. 3.7 can be followed. In this instance the fed-back residual would be treated as altering the limit cycle DIDF only, leading to a second-approximation limit cycle DIDF. In the case of analytic nonlinearities a second-order approximation can be obtained by perturbing the original DF solution and satisfying perturbation-equation first-order terms derived from the system differential equation. Finally, for limit cycling relay control systems with either command inputs which result in error-signal bias terms or asymmetric relays or both, Tsypkin's method, presented in Sec. 3.8, can be modified and otherwise directly extended to yield exact solutions.

REFERENCES

1. Atherton, D. P., G. F. Turnbull, A. Gelb, and W. E. Vander Velde: Discussion of the Double Input Describing Function (DIDF) for Unrelated Sinusoidal Signals, *IEEE Trans. Autom. Control*, vol. AC-9, no. 2 (April, 1964), pp. 197–198.
2. Furman, G. G.: Removing the Noise from the Quantization Process by Dithering: Linearization, *RAND Mem.* RM-3271-PR (February, 1963), pp. 1–40.
3. Gelb, A.: "The Analysis and Design of Limit Cycling Adaptive Automatic Control Systems," Sc.D. thesis, Massachusetts Institute of Technology, Cambridge, Mass., August, 1961.
4. Gelb, A.: The Foundation for Limit Cycling Adaptive Design, *Proc. Northeast Electronics Res. Eng. Meeting* (November, 1961), pp. 76–77.
5. Gelb, A.: The Dynamic Input-Output Analysis of Limit Cycling Control Systems, *Proc. JACC, New York University*, New York (June, 1962), pp. 9.3-1–9.3-11.
6. Gelb, A., and T. C. Blaschke: Design for Adaptive Roll Control of a Missile, *J. Aeron. Astronautics*, vol. 9, no. 4 (Winter Issue, 1962), pp. 99–105.
7. Gelb, A., and W. E. Vander Velde: On Limit Cycling Control Systems, *IEEE Trans. Autom. Control*, vol. AC-8, no. 2 (April, 1963), pp. 142–157.
8. Korolev, N. A.: Pulse Stabilization of Automatic Control Relay Systems, *Automation and Remote Control*, vol. 18, no. 5 (April, 1958), pp. 435–446.
9. Li, Y. T., and W. E. Vander Velde: Philosophy of Nonlinear Adaptive Systems, *Proc. First IFAC Congr.*, Moscow, U.S.S.R., 1960, Butterworth Scientific Publications, London.
10. Loeb, J. M.: A General Linearizing Process for Nonlinear Control Systems, *Manual and Autom. Control*, 1952, pp. 274–284, Butterworth Scientific Publications, London.
11. Lozier, J. C.: Carrier-controlled Relay Servos, *Elec. Eng.*, vol. 69 (December, 1950), pp. 1052–1056.
12. MacColl, L. A.: "Fundamental Theory of Servomechanisms," D. Van Nostrand Company, Inc., Princeton, N.J., 1945.
13. Minorsky, N.: On Asynchronous Action, *J. Franklin Inst.*, vol. 59 (March, 1955), pp. 209–219.
14. Mishkin, E., and L. Braun, Jr.: "Adaptive Control Systems," McGraw-Hill Book Company, New York, 1961.
15. Naslin, P.: A Simplified Theory of Feedback Control Systems, Part 12, *Process Control and Automation*, vol. 6, no. 6 (June, 1959), pp. 273–277.
16. Oldenburger, R.: Signal Stabilization of a Control System, *Trans. ASME*, vol. 79 (August, 1957), pp. 1869–1872.
17. Oldenburger, R., and C. C. Liu: Signal Stabilization of a Control System, *Trans. AIEE*, vol. 78 (May, 1959), pp. 96–100.
18. Oldenburger, R., and T. Nakada: Signal Stabilization of Self-oscillating Systems, *IRE Trans. Autom. Control*, vol. AC-6, no. 3 (September, 1961), pp. 319–325.
19. Oldenburger, R., and R. C. Boyer: Effects of Extra Sinusoidal Inputs to Nonlinear Systems, *Trans. ASME*, ser. D, *J. Basic Eng.*, vol. 84, no. 4 (December, 1962), pp. 559–570.
20. Popov, E. P.: "The Dynamics of Automatic Control Systems," Addison-Wesley Publishing Company, Inc., Reading, Mass., 1962.
21. Popov, E. P., and N. P. Pal'tov: "Priblizhennye metody issledovania nelineinych avtomaticheskikh sistem," Gosudarstvennoe Izdatelstvo Phizikio, Mathematicheskoi Literatur, Moscow, 1960. English translation: "Approximate Methods for Analyzing Non-linear Automatic Systems," Wright Patterson Air Force Base, Ohio, Translation Services Branch, Foreign Technology Division, January, 1963.

22. Shuck, O. H.: Honeywell's History and Philosophy in the Adaptive Control Field, *Wright Air Develop. Center, Tech. Rept.*, January, 1959, Wright Patterson Air Force Base, Ohio.

PROBLEMS

6-1. Show that any single-valued *asymmetric* nonlinearity can be represented as the parallel combination of one odd $[y(x) = -y(-x)]$ and one even $[y(x) = y(-x)]$ *symmetric* nonlinearity. What are the corresponding odd and even elements which comprise the two-segment piecewise-linear asymmetric nonlinearity

$$y(x) = \begin{cases} m_1 x & x \geq 0 \\ m_2 x & x < 0 \end{cases}$$

Calculate the DIDFs for these odd and even elements and sum the results, thus arriving at Eqs. (6.2-11) and (6.2-14).

6-2. (*a*) Show that the incremental-input describing function for a limiter of input breakpoints $\pm\delta$ and output saturation levels $\pm D$ is given by

$$N_B(A) = \frac{2D}{\pi\delta} \sin^{-1} \frac{\delta}{A}$$

Perform this calculation twice, once by taking the limit of $N_B(A,B)$ as $B \to 0$, and once by applying Eq. (6.1-9) directly.

(*b*) Compute the incremental-input describing function for an ideal relay directly in terms of Eq. (6.1-9). Note that the weights (i.e., strengths) of the impulse functions in y' are not unity.

6-3. Compute the limit cycle and signal DIDFs for an asymmetric polynomial nonlinearity described by

$$y(x) = \begin{cases} cx^2 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

6-4. Compute signal and limit cycle DIDF expressions for the energy-storing nonlinear element described by

$$y(x,\dot{x}) = x(\dot{x})^2$$

6-5. Show that N_B for an odd square-law characteristic is given by

$$N_B = \frac{A^2}{\pi B} \left\{ \left[1 + 2 \left(\frac{B}{A} \right)^2 \right] \sin^{-1} \left(\frac{B}{A} \right) + 3 \frac{B}{A} \sqrt{1 - \left(\frac{B}{A} \right)^2} \right\}$$

and hence derive the result that for small B/A the DF of the equivalent nonlinear element is $1.27A$.

6-6. Show that when n is an even integer, the DIDFs of Eqs. (6.2-21) and (6.2-23) are in error by a factor of the order of $(B/A)^{2n}$. In particular, for $n = 2$, show that $N_A(A,B)$ is in error by the amount $(A/3\pi)(B/A)^4$.

6-7. A frequently used measure of the dependence of closed-loop transmission F_{cl} upon an environmental parameter ξ is the sensitivity function $S_\xi^{F_{cl}}$. It is defined as the ratio of per-unit change in F_{cl} to per-unit change in ξ, viz.,

$$S_\xi^{F_{cl}} = \frac{dF_{cl}/F_{cl}}{d\xi/\xi}$$

Now consider two unity feedback systems with third-order linear elements $L(s)$,

one with an additional linear gain K, and one with a nonlinearity N. Assume $L(s)$ to depend in some way upon ξ. Let F_{ol} denote the open-loop transfer function.

(*a*) For the linear system show that

$$S_\xi^{F_{cl}} = \frac{1}{1 + F_{ol}} S_\xi^L$$

(*b*) For the nonlinear system show that

$$S_\xi^{F_{cl}} = \frac{1}{1 + F_{ol}} (S_\xi^L + S_\xi^{N_B})$$

(*c*) If ξ affects only the gain constant of $L(s)$, and the nonlinearity is an ideal relay ($N_B \approx 2D/\pi A$), show that $S_\xi^{F_{cl}} = 0$. (*Hint:* A useful relationship is $S_\xi^{F_{cl}} = S_\xi^{F_{ol}} S_{F_{ol}}^{F_{cl}}$).

6-8. (*a*) Find the complete description of the equivalent nonlinear element describing a relay with dead zone subjected to triangular-wave dither.

(*b*) Repeat (*a*) for a limiter with square-wave dither.

6-9. Find the equivalent nonlinear characteristic corresponding to an ideal relay with a sawtooth dither input, shown in Fig. 6-1.

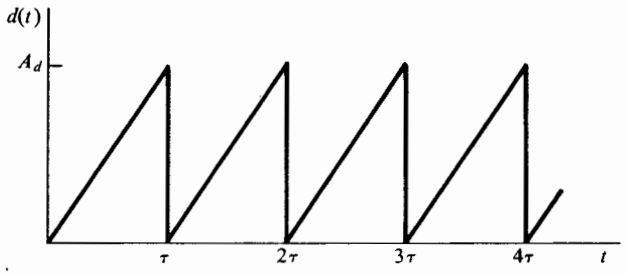


Figure 6-1 *Sawtooth dither waveform.*

6-10. The voltage induced in the antenna of a *radio compass* varies with the antenna orientation with respect to the transmitter, as shown in Fig. 6-2. Devise and analyze a system for making the radio compass direction sensitive (Naslin, *op. cit.*).

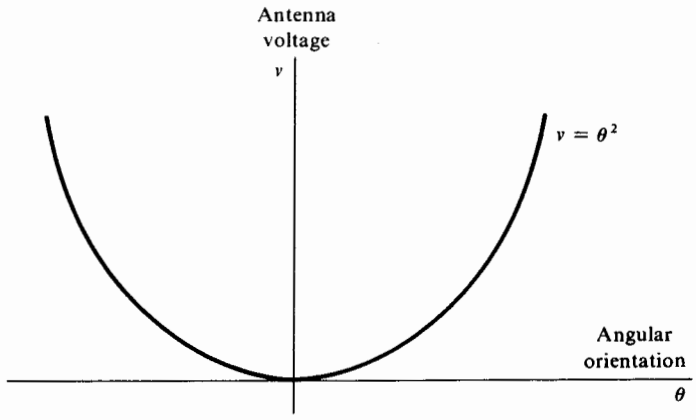


Figure 6-2 *Even symmetric antenna voltage-orientation characteristic.*

6-11. Determine the three nonlinear simultaneous equations which describe limit cycle behavior of the system shown in Fig. 6-3. Discuss an approach to the graphical solution of these equations. Plot the normalized quantities, A/δ, B/δ, $\omega_0\tau$ versus $DK\tau/\delta$.

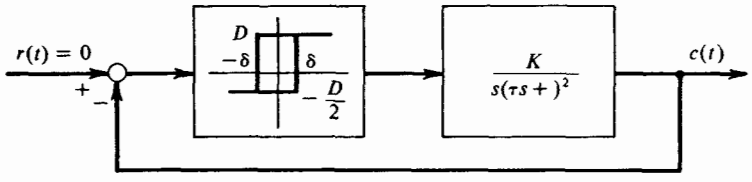


Figure 6-3 *System with an output-biased rectangular hysteresis nonlinearity.*

6-12. Find the limit cycle state of the system of Fig. 6-4 containing an asymmetric relay characteristic, if

(*a*) $$L(s) = \frac{sK}{(\tau s + 1)^4}$$

(*b*) $$L(s) = \frac{K}{(\tau s + 1)^3}$$

(*c*) $$L(s) = \frac{K}{s(\tau s + 1)^2}$$

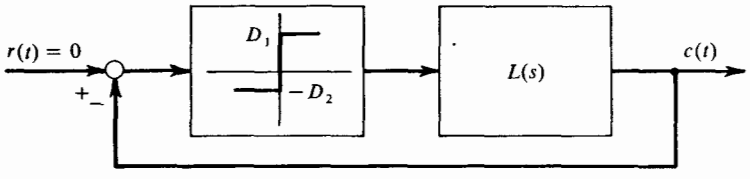


Figure 6-4 *System with an asymmetric relay nonlinearity.*

6-13. Describe the limit cycle state of a system containing an input-biased ideal relay with symmetric drive levels, described by

$$y(x) = \begin{cases} D & x > \delta \\ 0 & x = \delta \\ -D & x < \delta \end{cases}$$

and the linear elements

$$L(s) = \frac{K}{s(\tau s + 1)^2}$$

6-14. Design a simple adaptive controller for the variable plant $L(s) = 10\tau/(\tau s + 1)^3$, where $0.1 < \tau < 10$. Draw and interpret the closed-loop frequency response curves for the resultant system in the extremes $\tau = 0.1$ and $\tau = 10$. Compute the ramp response of the adaptive system. Design a simple gain-compensated linear closed-loop system for $L(s)$, and compare this system to the adaptive system.

6-15. The system of Fig. 6-5 has a square-law error detector. It can be made to respond to input signals by biasing the input signal off zero. Plot the bias at the output $c(t)$, and the input-output frequency response function for all significant frequencies with $B = 5$ units and $r(t) = 5 \sin \omega t$. What would be the effect of a negative bias?

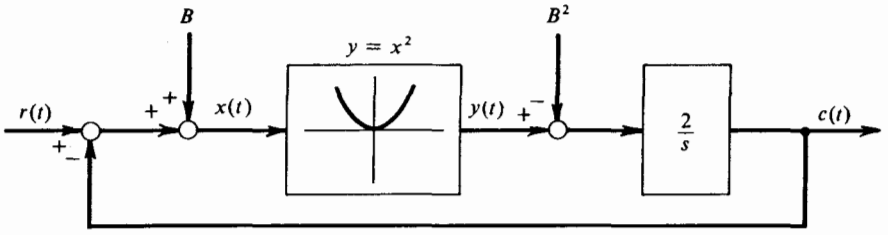


Figure 6-5

6-16. Consider the system of Fig. 6-6. Determine the limit cycle frequency and amplitude with the describing function method for $\tau_1 = 0$. Select a τ_1 to reduce the limit cycle amplitude to one-half of the uncompensated value, and determine the average steady-state error when the compensated system is driven by a ramp function $r(t) = 5t$.

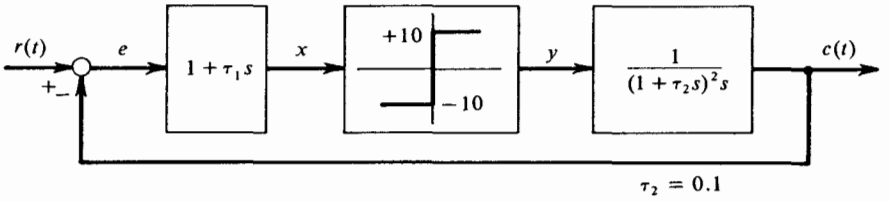


Figure 6-6

6-17. Design compensation for the inverted pendulum-balancing system shown in Fig. 6-7. The specifications are that the amplitude of the limit cycle at θ must be no greater than 5° and the error between θ and θ_r must be no greater than 1° for all constant $|\theta_r| \leq 30°$. A current switch is controlled by the sign of $x(t)$. The steady-state current produces a torque $M = 10$ lb-ft, but there is a first-order lag ($\tau = 0.1$ sec) in the current, and hence torque, buildup. The pendulum is essentially ideal, with length 1 ft and weight 10 lb. It is to be controlled to an angle off vertical (inverted) given by θ_r.

For your design, plot the average error, the limit cycle amplitude and frequency, and some measure of the dominant system response characteristics as functions of θ_r for $0 \leq \theta_r \leq 30°$.

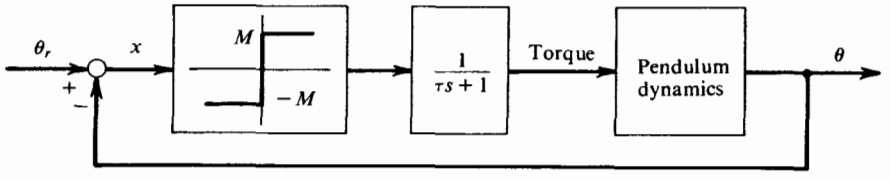


Figure 6-7

6-18. The laser gyroscope exhibits an input-rate (ω_i) output-frequency difference (Δf) characteristic which displays a dead zone due to "lock-in" (frequency entrainment), which arises from coupling between the two counterrotating laser beams. One method of enabling use of this gyro at low input rates is to apply a mechanical input-rate dither, to linearize the dead zone.

Assume that the nonlinear laser characteristic is described by

$$\Delta f = K\omega_i \qquad \omega_i < -\delta,\ \omega_i > \delta$$
$$= 0 \qquad -\delta \leq \omega_i \leq \delta$$

Describe the equivalent gyro characteristic resulting from an applied dither signal of the form

$$d(t) = A_d \sin \omega_d t \qquad A_d > \delta$$