THE STABILITY OF LIMIT CYCLES IN TIME-LAG RELAY CONTROL SYSTEMS

A Thesis

Presented to

Faculty of Graduate Studies and Research

of

The University of Manitoba

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Electrical Engineering

---

by

Barry Edward Brookes

May 1967

ABSTRACT

The failure of Loeb's Rule in a particular example leads to a general limit cycle stability analysis for time-lag relay control systems and ultimately, to the reason for this failure.

ACKNOWLEDGEMENT

The author thanks Professor R. A. Johnson for suggesting this topic and the National Research Council of Canada for supporting this work through Grant-in-Aid A584.

CONTENTS

LIST OF FIGURES

FIGURES PAGE

CHAPTER I

INTRODUCTION

The describing function, supplemented by Loeb's Criterion,[1] is a powerful, simple, and widely used stability analysis for non-linear feedback control systems. In this technique, the describing function locates possible limit cycles and Loeb's Criterion (or Loeb's Rule as it is often called) indicates whether each limit cycle is stable or unstable. This latter information is essential because only a stable limit cycle will appear as a physical oscillation. Thus, Loeb's Criterion plays a crucial role in the analysis.

Unfortunately, both the describing function method and Loeb's Criterion are approximate and in both cases the error is very difficult to estimate. Nevertheless, Loeb's Criterion does give correct results in a large number of practical cases. However, for at least one control system, Loeb's Criterion gives erroneous results.

1. THE PROBLEM

Origin of the Problem

The genesis of the problem was the failure of Loeb's Criterion for the system in Figure 1.

---

[1]J. Loeb, "Phénomènes Héréditaires dans les Servomécanismes; un Critérium Général de Stabilité", Annales des Télécommunications, 6(12): 346-356(1951).

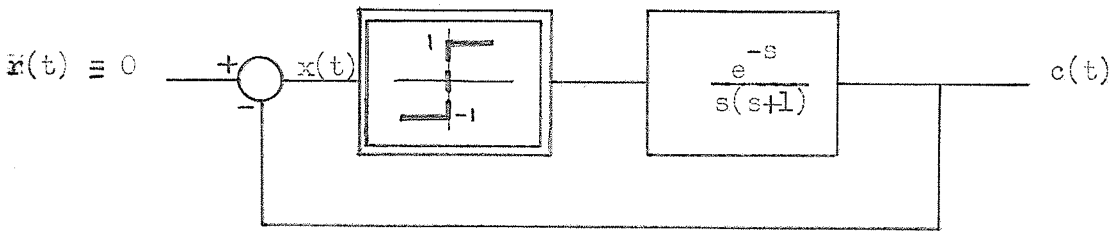


FIGURE 1

A SYSTEM FOR WHICH LOEB'S CRITERION FAILS

The conventional describing function analysis, in which the critical locus is superimposed on the Nyquist Diagram, is shown in Figure 2, for this system.[2]

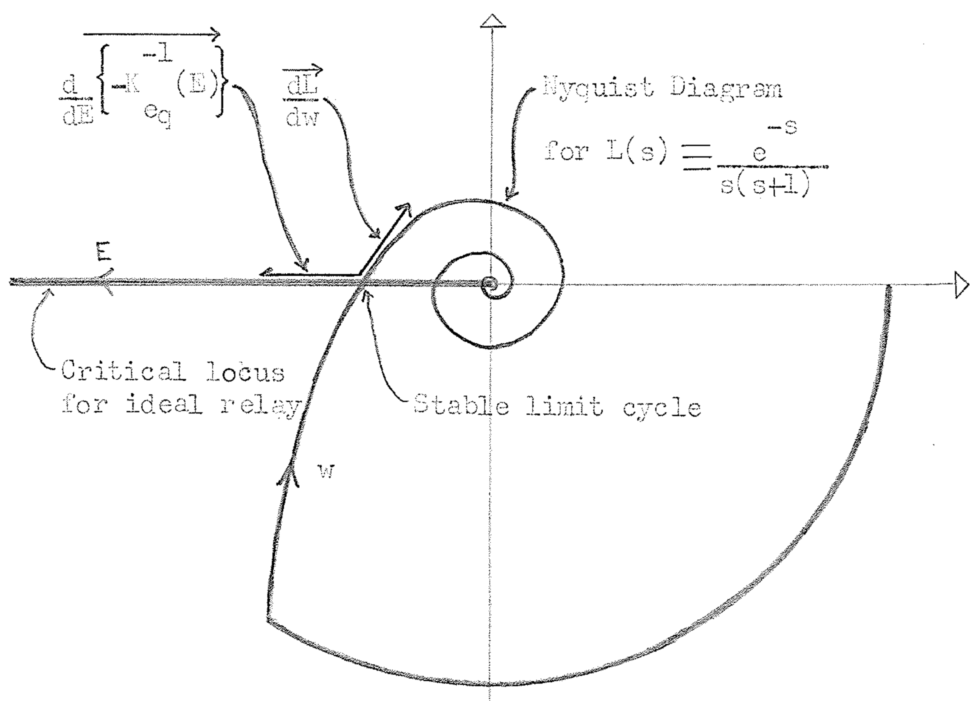


FIGURE 2

DESCRIBING FUNCTION - LOEB ANALYSIS FOR THE

SYSTEM SHOWN IN FIGURE 1

---

[2] The symbolism is that of John E. Gibson, Nonlinear Automatic Control, pp. 344-50.

To each intersection of the Nyquist Diagram with the critical locus, there corresponds a limit cycle whose frequency is that of the Nyquist diagram at the intersection and whose amplitude (E) may be read from the scaled critical locus at the intersection. Loeb's Criterion predicts a stable limit cycle if the vectorial cross product, denoted symbolically by $\overrightarrow{\frac{dL}{dw}}(jw) \times \overrightarrow{\frac{d}{dE}\left\{-K_{eq}^{-1}(E)\right\}}$, is positive at the intersection; an unstable limit cycle if the product is negative.[3] Therefore Loeb's Criterion predicts that all limit cycles in this system (Figure 1) are stable. This is a flagrant contradiction of the experimental fact that only the lowest frequency oscillation (which corresponds to the first intersection, counting from left to right in Figure 2) is stable.[4]

Statement of Thesis Objectives

A major objective of this thesis is to explain why Loeb's Criterion fails in the example just cited. But, the general discussion on limit cycle stability in Chapter II and the exact stability analysis of Chapter III, both of which are required to resolve this problem are each as significant as the explanation itself.

Significance of the Study

The failure of Loeb's Criterion in the previous example casts doubt upon its validity in all control systems. Moreover, by exposing the reason for this failure, this study further undermines confidence in the Criterion in general and in the positive vector product in particular. Furthermore it

---

[3] J. C. Gille, M. J. Pelegrin, and P. Decaulne, Feedback Control Systems, pp. 419-21.

[4] This was established independently on the analogue computer by two summer students at the University of Manitoba: Temple in 1965 and Carson in 1966.

shows that a reliable alternative analysis to the describing function - Loeb technique exists, at least for a restricted class of systems.

2. DEFINITIONS

The following definitions are deemed sufficiently precise for the purposes of this thesis:

Limit Cycles

An isolated periodic oscillation in a system will be called a limit cycle.[5] All periodic oscillations discussed herein are isolated; therefore they are all limit cycles.

Stability of Limit Cycles

A system will possess a stable (unstable) limit cycle if the linear variational equation about the limit cycle is stable (unstable). This definition of a stable limit cycle is similar, though not as mathematically rigorous, as that given by Hayashi for orbital stability of a trajectory.[6]

3. REVIEW OF THE LITERATURE

Loeb first propounded his criterion for testing the stability of limit cycles in 1951.[7] Gille et al. formulated this criterion in terms of the convenient cross product rule stated earlier.[8] In a later paper, Loeb developed a more complicated test which in first approximation reduced to the vector rule, but even this more complicated rule

---

[5] N. Minorsky, Nonlinear Oscillations, p. 71.

[6] C. Hayashi, Nonlinear Oscillations in Physical Systems, pp. 70-71.

[7] Loeb, loc. cit.

[8] Gille et al., loc. cit.

was not exact.[9] Furthermore, by assuming a differential equation throughout (see Equation (7) therein) he excluded time-lag systems (because they require a differential-difference equation for their representation). Grensted, in three interesting papers, developed a stability test for limit cycles using the operational calculus.[10] His work, however, suffered from the same shortcomings as Loeb's, namely it was approximate and it excluded time-lag systems.[11] This thesis surmounts these difficulties, but at present it is restricted to ideal relay control systems.

4. PLAN OF THE THESIS

First of all, by definition a limit cycle is stable or unstable accordingly as the linear variational equation about the limit cycle is stable or unstable. Consequently, in Chapter II a method is developed to derive this variational equation in a general single loop time-lag nonlinear feedback control system. In Chapter III, the specialization of this equation to the ideal relay system and the subsequent deduction of its stability properties brings the thesis objectives close to consummation. Indeed, when this knowledge is coupled with the extended Hamel

---

[9] J. Loeb, "Recent Advances in Nonlinear Servo Theory", (1953) In Frequency Response ed. R. Oldenburger, pp.260-67.

[10] P. E. W. Grensted, "The Frequency Response Analysis of Non-Linear Systems," Proc. I.E.E., Vol. 102, part C, 1955, pp.244-55.

P. E. W. Grensted, "Analysis of the Transient Response of Non-Linear Control Systems," A.S.M.E. Trans., Vol. 80, 1958, pp.427-32.

P. E. W. Grensted, "Frequency Response Methods Applied to Non-Linear Systems" In Progress in Control Engineering, Vol. 1, pp. 105-139.

[11] By assuming a rational operator form throughout, Grensted effectively excluded time-lag.

Locus the reason for the failure of Loeb's Rule is readily uncovered. In the fourth and final chapter, the work is summarized; conclusions are stated, and areas of further research are outlined.

CHAPTER II

THE LINEAR VARIATIONAL EQUATION

In order to determine the stability of a limit cycle it is necessary to find the linear variation (or equivalently the first variation) from the limit cycle. This may be found by substitution of the periodic solution plus a small perturbation (of unspecified form) in the system equation(s) in the traditional way. After the periodic solution is cancelled and all higher order terms in the variation are discarded the linear variational equation(s) remains (remain). Henceforth this (these) equation(s) will simply be referred to as the variational equation(s). Now if the system is specified by N state equations there will be N variational equations also. However, the presence of time-lag implies that the state (and hence the variational) equations are infinite in number.[1] Therefore, a more convenient method for deriving the variational equation for time-lag systems is formulated in this chapter. In particular, the explicit form of the variational equation for a saturating system is deduced.

Next, a connection between the state approach and operational approach for getting the variational equation is discussed for the zero

---

[1] R. A. Johnson, "State Space and Systems Incorporating Delay," Electronics Letters, Vol. 2, No. 7, July 1966, pp. 277-78.

time-lag case. This is followed by some remarks on the inadequacies of existing stability theorems.

For completeness, the nonautonomous case is also discussed, although the results are not used in the remainder of the thesis.

1. THE OPERATIONAL APPROACH

Cosgriff has given the following convenient method for finding the variational equation for the general system shown in Figure 3.[2]

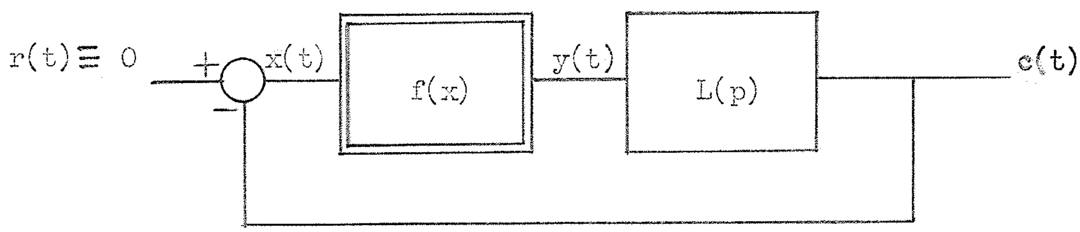


FIGURE 3

THE GENERAL SINGLE LOOP NONLINEAR FEEDBACK CONTROL SYSTEM

The operational differential equation for the autonomous system (Figure 3) is given by

$$L(p)\{f(x)\} + x = 0 \quad \text{.......... 1.}$$

If x' is a periodic solution of Equation (1) then the substitution of the perturbed solution, $x' + u$, into Equation (1) yields

$$L(p)\{f(x'+u)\} + x' + u = 0 \quad \text{.......... 2}$$

---

2 R. L. Cosgriff, "Application of Linear Differential Equations with Periodic Coefficients in the Study of Non-linear Phenomena", Proc. of First International Congress of the I.F.A.C., Vol. 2, 1960, pp. 883-87.

which, if $f(x)$ has **a Taylor series** with respect to x, becomes

$$L(p)\left\{f(x') + u\frac{df}{dx}(x') + \frac{u^2}{2!}\frac{d^2}{dx^2}f(x') + \cdots\right\} + x' + u = 0 \quad \ldots\ldots 3.$$

When the limit cycle solution,

$$L(p)\left\{f(x')\right\} + x' = 0,$$

is cancelled and the higher order terms in the perturbation, u, are discarded Equation (3) reduces to

$$L(p)\left\{u\frac{df}{dx}(x')\right\} + u = 0 \quad \ldots\ldots 4.$$

Since $x'(t)$ is a periodic function of time it follows that $\frac{df}{dx}(x')$ is also periodic and of the same period. Therefore Equation (4) is a linear differential equation with (time) periodic coefficients. Thus the nonlinear problem of limit cycle stability has been reduced to an equivalent (time-variable) linear problem. This is not surprising in view of the fact that only linear variations from the periodic solution were considered.

Cosgriff then proceeded to deduce $u(t)$ by setting $x'(t+\Delta t)$ equal to $x'(t)+u(t)$. This, however, is a questionable tactic since $x'(t+\Delta t)$ represents a perturbation along the limit cycle whereas $x'(t)+u(t)$ is a perturbation off the limit cycle. If these two perturbed solutions are to be identically equal for all time, then, contrary to the original hypothesis, $u(t)$ cannot be a perturbation off the limit cycle. The correct way to find either $u(t)$ or any of its properties is by a consideration of the variational equation (Equation (4)).

The method for finding the variational equation given by Cosgriff may easily be extended to the general nonlinear system with time-lag by replacement of $L(p)$ in Figure 3 by $G(p)\ e^{-\tau p}$, in which $G(p)$ is a rational, stable, and minimum phase function in p and $e^{-\tau p}$ represents a pure time-delay of τ seconds. For the autonomous case, the delay may be either in the forward or in the feedback path; the analysis is the same in both cases. Substitution of $G(p)\ e^{-\tau p}$ for $L(p)$ in Equation (4), yields

$$G(p)\ e^{-\tau p}\left(u\frac{df}{dx}(x')\right) + u = 0 \qquad (5)$$

which is the explicit form of the variational equation for the time-lag case. This equation, which is basic for later work, is equivalent to the system in Figure 4 below.

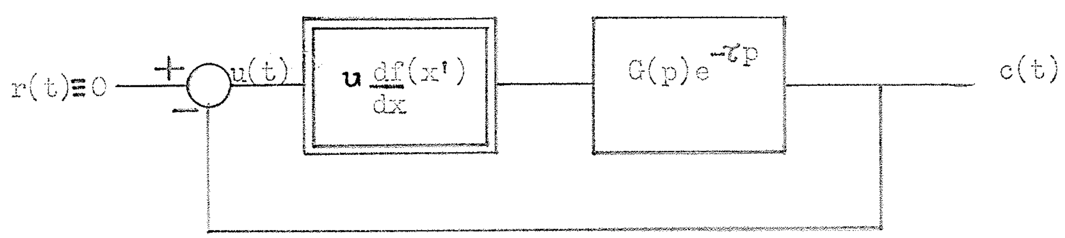


FIGURE 4

VARIATIONAL SYSTEM FOR THE GENERAL TIME-LAG NONLINEAR SYSTEM

Notes: 1. The function, $\frac{df}{dx}(x')$, is essentially a time variable gain factor (it is periodic in time with the same period as the limit cycle). Taken as a whole, this representation is a time-variable linear feedback control system with time-lag.

2. The symbols, $r(t)$ and $c(t)$, in this figure are not related to those in the system representation in Figure 3.

A lemma which will be useful later is that $\dot{x}'$ is always a solution to the variational equation (Equation (4)).[3] This may be established by differentiating the limit cycle solution,

$$L(p)\left\{f(x')\right\} + x' = 0,$$

with respect to time; the result (which proves the lemma) is Equation (4).

The advantage of this operational method is that a single variational equation is obtained rather than infinitely many as in the state approach. This method applies to all systems which can be represented by a single differential or differential-difference equation and, in particular, to all single-loop time-lag nonlinear feedback control systems (see Figure 3).

The Variational Equation for Time-Lag Systems With A Saturation Nonlinearity

The variational equation and its system representation will now be derived for the saturating system illustrated in Figure 5.

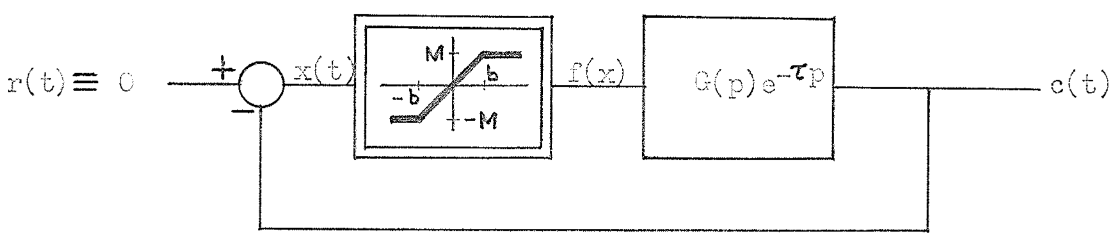


FIGURE 5

THE TIME-LAG SATURATING SYSTEM

Unfortunately, f(x) does not possess continuous derivatives when |x| is equal

---

3 E. A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, p. 322.

to b. Therefore $f(x'+u)$ cannot be expanded in a valid Taylor series at these points. Hence the following direct analysis is employed.

The saturation function is defined analytically by the relations:

$$f(x) = \frac{M}{b}x \quad \text{when } |x| < b \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 6a$$

and

$$f(x) = M\,\mathrm{sgn}(x) \quad \text{when } |x| \geq b \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 6b$$

which on perturbation become

$$f(x+u) = \frac{M}{b}(x+u) \quad \text{when } |x+u| < b \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 7a$$

and

$$f(x+u) = M\,\mathrm{sgn}(x+u) \text{ when } |x+u| \geq b \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 7b.$$

Since $u(t)$ may be made arbitrarily small initially, $|x+b|$ will differ from $|x|$ by less than any preassigned number for any finite b. Therefore, the perturbed solution at the saturation output may be written as

$$f(x'+u) = f(x') + \frac{uM}{b}\left\{U(x'+b) - U(x'-b)\right\} \quad \ldots\ldots\ldots\ldots \quad 8$$

in which $U(x)$ is the unit step function.$^{+}$ Now, $\frac{df}{dx}(x')$ is not defined at x' equal to $\pm b$. However, if the derivative is assigned the value zero at these points (without loss of generality) then Equation (8) may be rewritten as

$$f(x'+u) = f(x') + u\frac{df}{dx}(x') \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 9,$$

a result which produces the standard variational equation given by Equation (5).

If $x'(t)$ has odd symmetry about its zero crossings and if these zeroes are uniformly spaced

$+$ $U(0) = 1$

then a graphical evaluation of $\frac{df}{dx}(x')$ is given in Figure 6.[4] The time origin has been selected so that $x'(-\tau)$ is zero to agree with later work on the Hamel Locus.

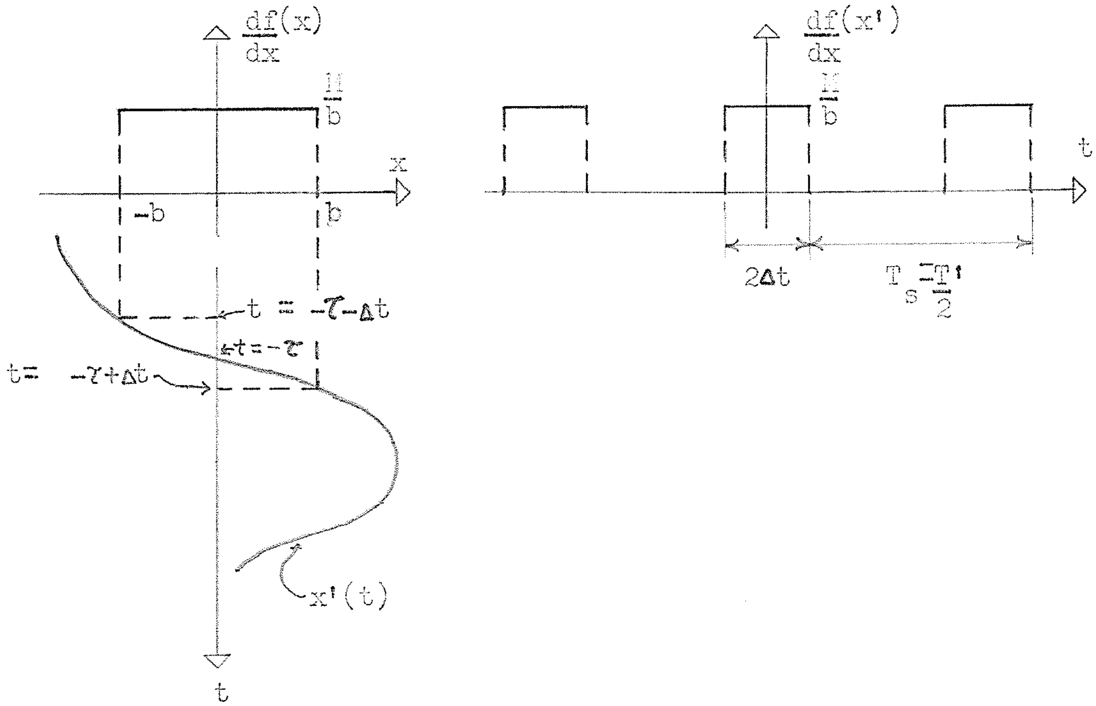


FIGURE 6

GRAPHICAL EVALUATION OF $\frac{df}{dx}(x')$ FOR THE SATURATING SYSTEM (FIGURE 5)

Definitions: T' = period of the limit cycle

T_s = period of $\frac{df}{dx}(x')$

Δt is defined such that:

$x'(-\tau-\Delta t) = -b$

The variational equation is equivalent to the non-ideal sampled data system in Figure 7. From Figure 6 it is evident that the sampling

[4] Similar results have been obtained for the zero time-lag case by Mohammed, "Steady-State Oscillations and Stability of On-off Feedback Systems," Ph.D. Thesis, U.B.C.

duration is $2\Delta t$ seconds and the sampling period, T_s, is exactly half the limit cycle period, T'.

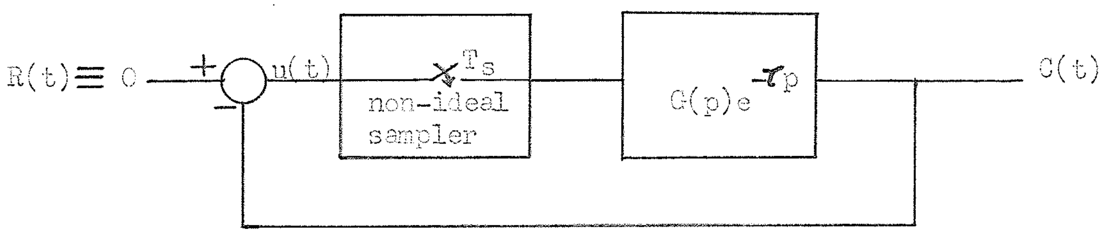


FIGURE 7

SYSTEM REPRESENTATION OF VARIATIONAL EQUATION

FOR SATURATING SYSTEM (FIGURE 5)

Note: The form of the non-ideal sampler is given by $\frac{df}{dx}(x')$ shown in Figure 6.

This system representation is fundamental for the results obtained in Chapter III.

2. THE AUTONOMOUS CASE WITHOUT TIME-LAG

This discussion is not in the main stream of the study since it deals with the case of zero time-lag. Nevertheless, it has merit because it establishes a link between the operational approach of Cosgriff and the traditional state equation approach used by mathematicians.

In the Appendix, the equivalence of these two methods is established for zero time-lag. The significance of this is that the large number of theorems which have been established by mathematicians for the state variational equations, now, by virtue of the Appendix, apply to the operational variational equation as well. Some of these theorems and their shortcomings are now reviewed.

The phase variational equations are, in general, a set of linear differential equations with (time) periodic coefficients. Such equations are difficult to solve though, in principle, they can be transformed to an equivalent set of linear constant coefficient equations by a Theorem of Liapunov.[5] The stability of a limit cycle is thus determined by a set of linear constant coefficient equations which can readily be solved in a variety of ways.

Using a different approach, Floquet gave the form of the general solution in terms of characteristic exponents with periodic coefficients.[6] In this formulation the stability of the limit cycle depends on the sign of the real part of the characteristic exponent. But no general method was given for finding these exponents and the reduction theorems available are of little use in determining the characteristic exponents, except in second order systems.

Both the Floquet and Liapunov methods are inconvenient for feedback systems because the system must first be converted to a state form. Furthermore, neither of these methods apply at all to the time-lag case because the appropriate equations are of the differential-difference variety.

---

[5] Pontryagin, L. S. Ordinary Differential Equations. Reading Mass.: Addison-Wesley Publishing Company, Inc., 1962, pp. 146-149.

[6] C. Hayashi. Nonlinear Oscillations in Physical Systems, pp. 82-86.

However, by the method of Chapter III, it is possible to transform the differential-difference variational equation of the time-lag relay control system to a stability equivalent linear time-invariant system and thereby, to circumvent the shortcomings of the previous theorems in a restricted class of time-lag systems. This transformation resembles Liapunov's in that the stability question of a linear time-variable system is reduced to a linear time-invariant problem. There are many systems, however, for which Liapunov's theorem applies when the method of Chapter III does not.

3. THE NONAUTONOMOUS CASE

The operational method is readily extended to nonautonomous systems by considering the driving function, $r(t)$, in Figure 3, not identically zero. The system differential (-difference) equation then becomes

$$L(p)\,(f\,(r-c)\,) - c(t) = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 10.$$

Limit cycle frequencies are, in general, multiples (superharmonics) or sub-multiples (subharmonics) of the input driving function $r(t)$.[7] Suppose $s(t)$ is the $\frac{n}{m}$ th sub (super) harmonic of $r(t)$ at the output then for all t the relation

$$s\,(t + \frac{n}{m} T_r\,) = s\,(t),$$

in which n, m are positive integers and T_r is the period of $r(t)$ is satisfied. Substitution of $s(t) + u(t)$ for $c(t)$ in Equation (10), and simplification of the result yields

---

[7] Combination tones are not considered here.

$$L(p)\left(u \frac{df}{dx}(r - s)\right) + u = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 11$$

as the variational equation for the nonautonomous case. This equation which is of the same basic form as the autonomous case is a linear differential (-difference) equation with time periodic coefficients.

The period of the coefficients is the period of $\frac{df}{dx}(r-s)$ which, in turn, is the period of $r(t) - s(t)$. If $s(t)$ is a subharmonic of $r(t)$ then the period of the coefficients is the period of $s(t)$, the limit cycle. On the other hand, if $s(t)$ is a super-harmonic of $r(t)$ then the coefficients assume the period of $r(t)$ (which is not the limit cycle period). Thus, the period of the coefficients in the nonautonomous case need not be the same as the limit cycle period.

CHAPTER III

STABILITY OF LIMIT CYCLES IN AUTONOMOUS TIME-LAG RELAY CONTROL SYSTEMS

In this chapter, the Hamel (and subsequently the Tsypkin) Locus for a system with an ideal relay is extended to include time-lag; this enables limit cycles therein to be determined exactly. The describing function could have been used for this purpose, but since it is an approximate method an intersection of the Nyquist Diagram with the critical locus does not guarantee a limit cycle. Thus, a meaningless situation could have arisen in which the stability of a false limit cycle was being investigated by an exact test. Such a situation, however, is precluded with an exact method like the Hamel Locus.

The next step is to specialize the variational equation for the saturating system developed in Chapter II to the case of an ideal relay. This equation is then solved by sampled data methods. This solution, together with the extended Hamel Locus, constitutes a complete and exact limit cycle and limit cycle stability analysis for the particular case of a time-lag relay control system. This technique is then used to analyse, in detail, the system which initiated the study. With the aid of this solution, an explanation is offered for the failure of Loeb's Rule.

The chapter concludes with a remark on the connection between Loeb's Criterion and a rule for the Tsypkin Locus, together with a rigorous derivation of the latter rule to supplant the fallacious existing one.

1. THE EXTENDED HAMEL LOCUS

This discussion, though similar to that of Gille et al., differs from the latter by the inclusion of a time-lag.[1] In particular, the time-lag system in Figure 8 is to be investigated for limit cycles in the autonomous state.

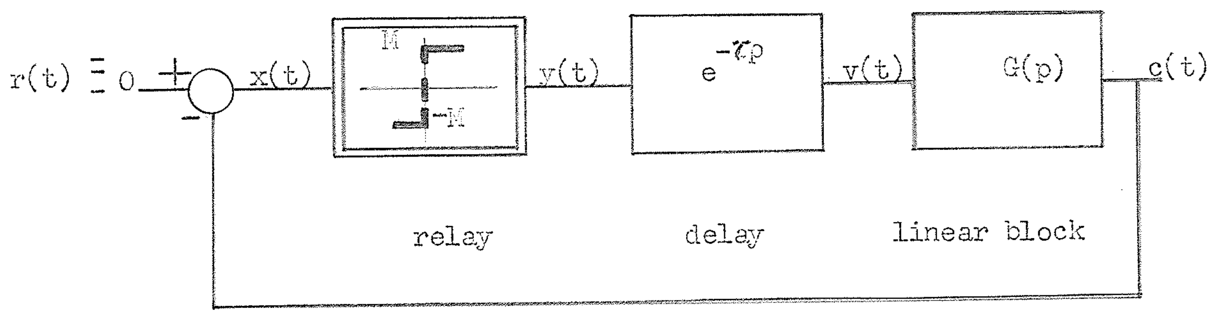


FIGURE 8

GENERAL TIME-LAG RELAY CONTROL SYSTEM

Note: G(s) is a stable minimum phase rational function of s, with simple non-zero poles.

If a symmetric periodic solution[2] of period T, with two switches per period,[3] is assumed to exist for x(t) and if the time origin is chosen

---

[1] Gille et al., op. cit., pp. 451-54.

[2] The nonlinearity is symmetric, so a periodic solution will be also.

[3] Oscillations with more than two switches per period are not considered here. Some work in this area has been done for the zero time-lag case by Mufti using a (state) matrix approach. See I. H. Mufti, "A Method For the Exact Determination of Periodic Motions in Relay Control Systems", Proc. of JACC, Troy N.Y., June 1965.

to agree with earlier work then v(t), as shown in Figure 9, will jump from $-M$ to $+M$ at the time origin. Consequently, a minus to plus commutation of the relay must have occurred τ seconds earlier, namely at t equal to $-\tau$. By the principle of causality, the output, c(t), at t equal to $-\tau$ is determined by the excitation, v(t), from t equals $-\infty$ to t equals $-\tau$. The explicit form of v(t) in this range depends on the length of delay (τ) relative to the limit cycle period length (T). In particular, a unique positive integer, m, may be defined for each value of $\frac{T}{2}$ such that

$$(m - 1)\frac{T}{2} < \tau \leqslant \frac{mT}{2} \quad \text{.................................} \quad 12a$$

or equivalently

$$\frac{\tau}{m-1} > \frac{T}{2} \geqslant \frac{\tau}{m} \quad \text{.................................} \quad 12b$$

is satisfied.

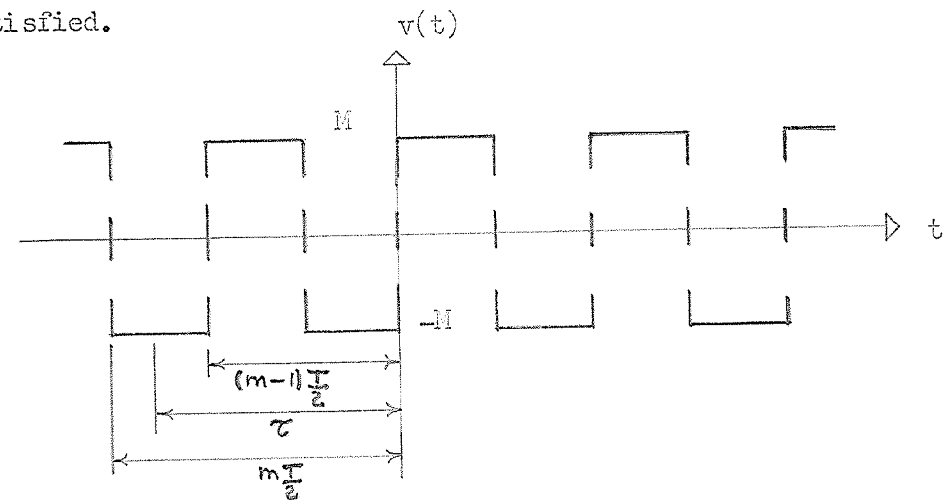


FIGURE 9

RESPONSE AT THE OUTPUT OF THE DELAY

The excitation, v(t), considered as a sequence of delayed step functions of height M, is given by

$$v(t) = M\left\{(-1)^m U(t+\frac{mT}{2}) + \sum_{k=m+1}^{\infty}(-1)^k \left[U(t+\frac{kT}{2}) - U(t+(k-1)\frac{T}{2})\right]\right\} \quad \text{......} \quad 13$$

in which U(t) is the unit step function. Substitution of the appropriately delayed unit step response of G(s), denoted by q(t+appropriate delay), for each delayed unit step function in Equation (13) produces

$$c(t) = M\left\{(-1)^m q(t+\frac{mT}{2}) + \sum_{k=m+1}^{\infty}(-1)^k \left[q(t+\frac{kT}{2}) - q(t+(k-1)\frac{T}{2})\right]\right\} \quad \text{......} \quad 14$$

for the response at the output of the system.

Now, G(s), being rational may be expressed in a partial fraction expansion as

$$G(s) = \sum_{i=1}^{N} \frac{K_i}{s-p_i} \quad \text{......} \quad 15$$

in which p_i is the i^{th} pole with residue K_i and N is the total number of poles of G(s). The unit step response of G(s) is determined from

$$q(t) = \mathcal{L}^{-1}\left\{\frac{G(s)}{s}\right\} = \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{G(s)}{s} e^{st} ds \quad \text{......} \quad 16$$

which on inversion yields

$$q(t) = \left\{-\sum_{i=1}^{N} \frac{K_i}{p_i} + \sum_{i=1}^{N} \frac{K_i}{p_i} e^{p_i t}\right\} U(t) \quad \text{......} \quad 17$$

in terms of the symbols of Equation (15). Substitution of this expression for q(t) in Equation (14) gives

$$c(t) = M\left\{(-1)^m \left[-\sum_{i=1}^{N} \frac{K_i}{p_i} + \sum_{i=1}^{N} \frac{K_i}{p_i} e^{p_i(t+\frac{mT}{2})}\right] + \sum_{i=1}^{N} \frac{K_i}{p_i} e^{p_i t}\left(1 - e^{-p_i\frac{T}{2}}\right) \sum_{k=m+1}^{\infty}(-1)^k e^{kp_i\frac{T}{2}}\right\}$$

$$\text{......} \quad 18$$

which, on summation of the geometric series,[4] reduces to

$$c(t) = M\left\{(-1)^m\left[-\sum_{i=1}^{N}\frac{K_i}{p_i} + \sum_{i=1}^{N}\frac{K_i}{p_i}e^{p_i(t+m\frac{T}{2})}\right] + \sum_{i=1}^{N}\frac{K_i e^{p_i t}\left[1 - e^{-p_i\frac{T}{2}}\right](-1)^{m+1}e^{p_i(m+1)\frac{T}{2}}}{p_i\left(1+e^{p_i\frac{T}{2}}\right)}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots \; 19$$

which simplifies to

$$c_m(t) = M(-1)^m\left\{-\sum_{i=1}^{N}\frac{K_i}{p_i} + 2\sum_{i=1}^{N}\frac{K_i e^{p_i(t+m\frac{T}{2})}}{p_i\left(1+e^{p_i\frac{T}{2}}\right)}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 20a$$

which on differentiation gives

$$\dot{c}_m(t) = 2M(-1)^m\sum_{i=1}^{N}\frac{K_i e^{p_i(t+m\frac{T}{2})}}{1+e^{p_i\frac{T}{2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 20b.$$

The subscript has been affixed to assert the dependence on m.

In analogy to the zero time-lag case, a Hamel Function, $H_m(\frac{T}{2})$ is defined on the range

$$\frac{\tau}{m-1} > \frac{T}{2} \geqslant \frac{\tau}{m}$$

as

$$H_m\left(\frac{T}{2}\right) = x_m\left(-\tau+\frac{T}{2}\right) + j\,\dot{x}_m\left(-\tau+\frac{T}{2}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 21.$$

---

[4] This geometric series whose common ratio is $e^{p_i\frac{T}{2}}$ converges since G(p) is a stable function (that is, Real$\{p_i\}$ is less than zero for all i).

Since the system is autonomous and the oscillation symmetric, the real and imaginary parts may be written as

$$\mathrm{Re}\left\{H_m\left(\frac{T}{2}\right)\right\} = x_m\left(-\tau+\frac{T}{2}\right) = -c_m\left(-\tau+\frac{T}{2}\right) = c_m(-\tau) \quad \ldots\ldots\ldots\ldots\ldots 22a$$

and

$$\mathrm{Im}\left\{H_m\left(\frac{T}{2}\right)\right\} = \dot{x}_m\left(-\tau+\frac{T}{2}\right) = -\dot{c}_m\left(-\tau+\frac{T}{2}\right) = \dot{c}_m(-\tau) \quad \ldots\ldots\ldots\ldots\ldots 22b.$$

Evaluation of Equations (20a, b) at t equals $-\tau$ and insertion of the result in Equations (22a, b) produces

$$\mathrm{Re}\left\{H_m\left(\frac{T}{2}\right)\right\} = M(-1)^m\left\{-\sum_{i=1}^{N}\frac{K_i}{p_i} + 2\sum_{i=1}^{N}\frac{K_i e^{p_i\left(-\tau+\frac{mT}{2}\right)}}{p_i\left(1+e^{\frac{p_iT}{2}}\right)}\right\} \quad \ldots\ldots\ldots\ldots\ldots 23a$$

and

$$\mathrm{Im}\left\{H_m\left(\frac{T}{2}\right)\right\} = 2M(-1)^m\sum_{i=1}^{N}\frac{K_i e^{p_i\left(-\tau+\frac{mT}{2}\right)}}{1+e^{\frac{p_iT}{2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 23b$$

which is an explicit form of the Hamel Function in terms of the partial fraction symbolism. The set of Hamel Functions covering all possible limit cycles, namely

$H_1\left(\frac{T}{2}\right)$ for $\infty > \frac{T}{2} \geq \tau$

$H_2\left(\frac{T}{2}\right)$ for $\tau > \frac{T}{2} \geq \frac{\tau}{2}$

$H_3\left(\frac{T}{2}\right)$ for $\frac{\tau}{2} > \frac{T}{2} \geq \frac{\tau}{3}$

$\vdots$

plotted in tandem in the complex plane with $\frac{T}{2}$ as the running parameter, comprises the Hamel Locus, $H\left(\frac{T}{2}\right)$. This formulation contains the usual (zero time-lag) Hamel Locus as a special case. For, if the time-lag is zero then m is unity and Equation (21) reduces to

$$H_1\left(\frac{T}{2}\right) = x_1\left(\frac{T}{2}\right) + j\, \dot{x}_1\left(\frac{T}{2}\right)$$

on the range

$$\infty > \frac{T}{2} \geqslant 0$$

which except for the subscript, is identical to the usual Hamel Locus.

The conditions for a limit cycle for an ideal relay are

$$\mathrm{Re}\left\{H_m\left(\frac{T}{2}\right)\right\} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 24a$$

and

$$\mathrm{Im}\left\{H_m\left(\frac{T}{2}\right)\right\} < 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 24b.$$

If for some value of T, say T_m, in the range

$$\frac{\tau}{m-1} > \frac{T}{2} \geqslant \frac{\tau}{m}$$

these conditions (Equations (24a, b)) are satisfied simultaneously, then the system will possess a limit cycle of period T_m.[5] Thus, every intersection of the Hamel Locus, $H\left(\frac{T}{2}\right)$, with the negative imaginary axis corresponds to a limit cycle.

The Hamel Function for the case in which G(p) contains a single integration may be deduced from Equations (23a, b) (without loss of generality) by taking the limit as p_N tends to zero. Evaluation of this limit gives

$$\mathrm{Re}\left\{H_m\left(\frac{T}{2}\right)\right\} = M(-1)^m \left\{ -\sum_{i=1}^{N-1} \frac{K_i}{p_i} + 2\sum_{i=1}^{N-1} \frac{K_i\, e^{p_i\left(-\tau + \frac{mT}{2}\right)}}{p_i\left(1 + e^{\frac{p_i T}{2}}\right)} + K_N\left[(2m-1)\frac{T}{4} - \tau\right] \right\} \quad \ldots 25a$$

---

[5] This does not mean that there will be a limit cycle for each value of m.

$$\operatorname{Im}\left\{H_m\left(\frac{T}{2}\right)\right\} = 2M(-1)^m\left\{\sum_{i=1}^{N-1}\frac{K_i\, e^{p_i\left(-\tau+m\frac{T}{2}\right)}}{1+e^{p_i\frac{T}{2}}} + K_N\right\} \qquad \text{25b}$$

a result which will be useful later.

<u>Continuity of the Hamel Locus</u>

The Hamel Locus for time-lag systems need not be continuous at the joining points of the Hamel Functions. Indeed, $H(\frac{T}{2})$ is continuous only if the unit step response of $G(p)$ and its derivative (that is, $q(t)$ and $\dot{q}(t)$) are continuous. This will be so if $G(p)$ falls off at least as fast as $\frac{1}{p^2}$ for large p.[6] Similarly, the derivative of $H(\frac{T}{2})$ with respect to T will be continuous only if $G(p)$ falls off at least as quickly as $\frac{1}{p^3}$ for large p.

2. THE EXTENDED TSYPKIN LOCUS

Traditionally, a Fourier Series analysis of the relay control system (Figure 8) led to the Tsypkin Locus, $\Lambda(w)$, given by

$$\Lambda(w) = \frac{4M}{\pi}\left\{\sum_{n=1,3,5\ldots}^{\infty} U(nw) + j\sum_{n=1,3,5\ldots}^{\infty}\frac{V(nw)}{n}\right\} \qquad 26$$

in which $U(nw)$ and $V(nw)$ are defined by

$$U(nw) \equiv \operatorname{Re}\left\{G(jnw)\,e^{-jnw\tau}\right\} = |G(jnw)|\cos(\angle G(jnw) - nw\tau) \qquad \text{27a}$$

and

$$V(nw) \equiv \operatorname{Im}\left\{G(jnw)\,e^{-jnw\tau}\right\} = |G(jnw)|\sin(\angle G(jnw) - nw\tau) \qquad \text{27b.}$$

For the ideal relay, the conditions for a limit cycle are

$$\operatorname{Im}\left\{\Lambda(w)\right\} = 0 \qquad \text{28a}$$

---

[6] This follows from the initial value theorem.

and

$$Re\left\{\Lambda(w)\right\} < 0 \quad \ldots\ldots 28b.$$

Thus, the frequencies at which $\Lambda(w)$ intersects the negative real axis are the frequencies of the limit cycles.

In the zero time-lag case, the sums in Equations (27a,b) are readily evaluated by a theorem from complex variables. In the time-lag case, however, the theorem does not apply. Moreover, there seems to be no simple way to evaluate these sums.

This impasse was circumvented by defining the Tsypkin Function, $\Lambda_m(\omega)$ (in analogy with the Hamel Function) in terms of the error signal at the switching points, as

$$\Lambda_m(w) \equiv \frac{1}{w}\dot{x}_m(-\tau+\frac{\pi}{w}) + jx_m(-\tau+\frac{\pi}{w}) \quad \ldots\ldots 29$$

on the range

$$(m-1)\frac{\pi}{\tau} < w \leq m\frac{\pi}{\tau}$$

in which m is the positive integer defined for the Hamel Function. The real and imaginary parts of the Tsypkin Function are related to those of the Hamel Function by

$$Re\left\{\Lambda_m(w)\right\} = \frac{1}{w}Im\left\{H_m(\frac{T}{2})\right\} \quad \ldots\ldots 30a$$

and

$$Im\left\{\Lambda_m(w)\right\} = Re\left\{H_m(\frac{T}{2})\right\} \quad \ldots\ldots 30b.$$

The substitution of Equations (23a, b), in these relations produces

$$\mathrm{Re}\,\Lambda_m(w) = \frac{2M}{w}(-1)^m \sum_{i=1}^{N} \frac{K_i e^{p_i(-\tau+\frac{mT}{2})}}{1+e^{\frac{p_i T}{2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 31a$$

and

$$\mathrm{Im}\,\Lambda_m(w) = M(-1)^m \left\{ -\sum_{i=1}^{N} \frac{K_i}{p_i} + 2\sum_{i=1}^{N} \frac{K_i e^{p_i(-\tau+\frac{mT}{2})}}{p_i(1+e^{\frac{p_i T}{2}})} \right\} \ldots\ldots\ldots\ldots\ldots 31b.$$

The set of Tsypkin functions:

$$\Lambda_1(w) \quad \text{for} \quad 0 < w \leqslant \frac{\pi}{\tau}$$
$$\Lambda_2(w) \quad \text{for} \quad \frac{\pi}{\tau} < w \leqslant \frac{2\pi}{\tau}$$
$$\Lambda_3(w) \quad \text{for} \quad \frac{\pi 2}{\tau} < w \leqslant \frac{3\pi}{\tau}$$
$$\vdots$$

plotted in tandem in the complex plane with w as the running parameter comprises the Tsypkin Locus, $\Lambda(w)$.

Relationship to the Describing Function Analysis

Since $G(p)$ is a rational function, at high frequencies $G(jw)$ falls off at least as fast as $\frac{1}{w}$. Therefore, $U(nw)$ and $V(nw)$ as given by Equations (27a, b), respectively, each tend to zero as w increases. Thus, at high frequencies, the Tsypkin Locus, defined by Equation (26), asymptotically approaches,

$$\Lambda(w) \simeq \frac{4M}{\pi}(U(w) + jV(w)) = \frac{4ML}{\pi}(jw) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 32,$$

since the higher order harmonics are negligible compared to the first. Whence, the Tsypkin Locus differs from the Nyquist Diagram by only a constant of proportionality at high frequencies. Furthermore, for an ideal relay, the critical locus is coincident with the Tsypkin commutation line, that is, with the negative real axis. Therefore, the intersection frequencies of the Tsypkin Locus at high frequencies are the same as those of the describing function analysis.

3. THE VARIATIONAL EQUATION FOR A TIME-LAG RELAY SYSTEM

The variational equation for a time-lag relay control system (Figure 8) is easily obtained as a limiting case of the saturating system dealt with in Chapter II. For convenience, the form of $\frac{df(x')}{dx}$ for saturation is repeated in Figure 10 below.

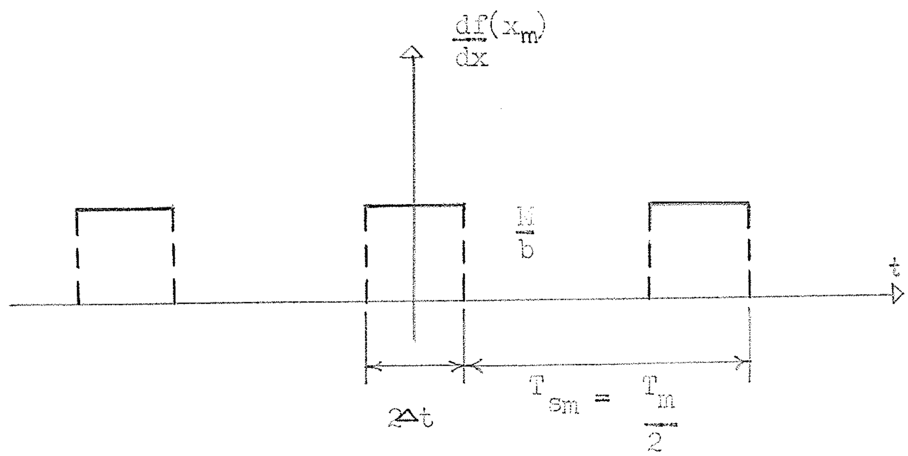


FIGURE 10

TIME RESPONSE OF NON-IDEAL SAMPLER FOR SATURATING VARIATIONAL SYSTEM

Definitions: T_m = period of the m^{th} limit cycle

T_{sm} = period of the sampler for the m^{th} variational system

$x_m(t)$ = mth limit cycle

Note: The subscript notation of the extended Hamel Locus will be used henceforth in preference to the prime notation of Chapter II. For example, T_m will be used to denote the limit cycle period rather than T'.

As Δt, b tend to zero the saturation nonlinearity approaches the ideal relay and $\frac{df}{dx}(x')$ tends toward a train of impulses each of weight, A_m, given by

$$A_m \doteq 2M \lim_{\Delta t, b \to 0} \left\{ \frac{\Delta t}{b} \right\} \qquad 33$$

which may be evaluated with the aid of

$$x_m(-\tau - \Delta t) \doteq x_m(-\tau) - \Delta t \dot{x}_m(-\tau) \qquad 34$$

a Taylor approximation which becomes exact in the limit. The first condition for oscillation (see Equation (24a)) requires that $x_m(-\tau)$ be zero. Furthermore, from Figure 6, Δt is defined by

$$x_m(-\tau - \Delta t) \doteq -b \qquad 35$$

for all b. Insertion of these results in Equation (34) produces

$$b = \Delta t \dot{x}_m(-\tau) \qquad 36$$

which, on rearrangement, becomes

$$\lim_{\Delta t, b \to 0} \left\{ \frac{\Delta t}{b} \right\} = \frac{1}{\dot{x}_m(-\tau)} \qquad 37$$

in the limit. Substitution of this limit in Equation (33) yields

$$A_m = \frac{2M}{\dot{x}_m(-\tau)} \qquad 38$$

a result which is in agreement with Gille et al.[7]

The variable gain function or sampling function, $\frac{df}{dx}(x_m)$, written

---

[7] His approach is quite different. Gille et al, *op. cit.*, pp. 478-80.

as an impulse train becomes

$$\frac{df}{dx}(x_m) = A_m \sum_{k=-\infty}^{\infty} \delta(t-kT_{sm}) \quad \text{.......................... } 39$$

in which $\delta(t)$ is the unit impulse function. Substitution of this expression for $\frac{df}{dx}(x)$ in Equation (5), gives

$$G(p)e^{-\tau p}\left\{u(t) A_m \sum_{k=-\infty}^{\infty} \delta(t-kT_{sm})\right\} + u(t) = 0 \quad \text{..............} 40$$

which in standard sampled-data notation[8] becomes

$$A_m G(p)e^{-\tau p}\left\{u_m^*(t)\right\} + u_m(t) = 0 \quad \text{......................} 41$$

when subscripts are affixed.

A System Representation

Equation (41) is represented by the conventional sampled-data system in Figure 11. The stability of the variational system is governed by the position of the zeroes of $H_m^*(s)$, given by

$$H_m^*(s) = 1 + A_m L_m^*(s) \quad \text{................................ } 42.$$

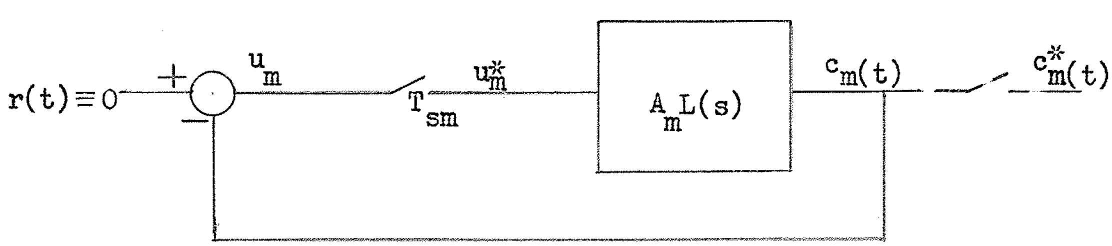


FIGURE 11

VARIATIONAL SYSTEM FOR RELAY CONTROL SYSTEM

---

8 Benjamin C. Kuo, Analysis and Synthesis of Sampled-Data Control Systems, Chapter 2.

Notes: 1. The variable s replaces p in this figure because a transform rather than an operator interpretation of the linear block is desired.

2. $L(s)$ equals $G(s)e^{-\tau s}$.

The variational system is stable if $H_m^*(s)$ has no zeroes in the right hand s-plane; it is unstable in the contrary case.

4. TRANSFORMATION OF THE VARIATIONAL SYSTEM

In this section, which closely parallels the work of Kuo, a linear time-invariant system is obtained that has the same stability properties as the variational system (Figure 11).[9] Thus, in the sense of stability, the variational system is transformed to an equivalent linear time-invariant system. The merit of this conversion is that the transformed system is amenable to standard stability techniques whereas the variational system is not.

Since the function of interest, $H_m^*(s)$ differs from $L_m^*(s)$ by only an invariant, it is convenient (and sufficient) to find the explicit form of the latter. To this end, let the inverse transform of $L(s)$ be defined by

$$l(t) = \mathcal{L}^{-1}\left\{L(s)\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 43$$

which on substitution of $G(s)e^{-\tau s}$ for $L(s)$ (Note 2, Figure 11) becomes

$$l(t) = \mathcal{L}^{-1}(G(s)e^{-\tau s}) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 44.$$

---

9 This discussion is more general than that of Kuo because it includes time-lag systems. Kuo, op. cit., pp. 73-6, 106-7, 157-63.

Insertion of the partial fraction expansion for G(s) from Equation (15), in Equation (44) gives

$$l(t) = \mathcal{L}^{-1}\left\{\sum_{i=1}^{N} \frac{K_i}{s-p_i} e^{-\tau s}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 45$$

which on inversion yields

$$l(t) = U(t-\tau) \sum_{i=1}^{N} K_i e^{p_i(t-\tau)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 46$$

in which U(t) is the unit step function.

When Equation (46) is sampled at the frequency of the variational system (Figure 11) it becomes

$$l_m^*(t) = U(t-\tau) \sum_{i=1}^{N} K_i e^{p_i(t-\tau)} \sum_{k=m}^{\infty} \delta(t-kT_{sm}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 47$$

the transform of which is given by

$$L_m^*(s) = \int_0^{\infty}\left\{e^{-st} U(t-\tau) \sum_{i=1}^{N} K_i e^{p_i(t-\tau)} \sum_{k=m}^{\infty} \delta(t-kT_{sm})\right\} dt \ldots\ldots\ldots\ldots \quad 48$$

which on evaluation reduces to

$$L_m^*(s) = \sum_{k=m}^{\infty} e^{-ksT_{sm}} \sum_{i=1}^{N} K_i e^{p_i(kT_{sm}-\tau)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 49$$

in which m, a positive integer, satisfies

$$m-1 < \frac{\tau}{T_{sm}} \leq m$$

as always. When the summation order is interchanged and the geometric series summed[10] the above equation simplifies to

---

[10] This series, with common ratio $e^{-T_{sm}(s-p_i)}$ converges since there exists an s for which $Re(s-p_i)$ is positive for all **i.**

$$L_m^*(s) = \sum_{i=1}^{N} \frac{K_i e^{-p_i \tau} e^{-mT_{sm}(s-p_i)}}{1 - e^{-T_{sm}(s-p_i)}} \quad \ldots\ldots\ldots\ldots \quad 50.$$

Note: In Equation (50), both m and T_{sm} are fixed numbers corresponding to a particular limit cycle of period T_m. In the Hamel Locus, however, both m and T are running parameters.

The form of $L_m^*(s)$ given by Equation (50), is somewhat intractable because of the presence of infinitely many poles. However, the number of poles becomes finite under the so-called z-transformation,

$$z = e^{sT_{sm}} \quad \ldots\ldots\ldots\ldots \quad 51,$$

by which the left-hand s-plane is mapped into the unit disc in the z-plane.[†]

Insertion of this transformation in Equation (50) and simplification of the result produces

$$L_m(z) = \frac{1}{z^{m-1}} \sum_{i=1}^{N} K_i e^{p_i(mT_{sm}-\tau)} \frac{1}{z - e^{p_i T_{sm}}} \quad \ldots\ldots\ldots\ldots \quad 53$$

in which

$$L_m(z) = L_m^*(s)\Big|_{s = \frac{1}{T_{sm}} \ln z}.$$

Although Equation (53) has only finitely many poles, conventional

---

[†] Equation (52) has been deleted.

[11] Deleted

stability techniques cannot be applied to it without considerable modification. However, under the r-transformation,[12]

$$z = \frac{r+1}{r-1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 54,$$

which maps the unit disc of the z-plane into the finite left-hand r-plane, Equation (53) becomes

$$L_m\left(\frac{r+1}{r-1}\right) = \frac{(r-1)^m}{(r+1)^{m-1}} \sum_{i=1}^{N} \frac{K_i e^{p_i(mT_{sm}-\tau)}}{(1-e^{p_iT_{sm}})r + 1 + e^{p_iT_{sm}}} \qquad \ldots\ldots\ldots\ldots \quad 55$$

which is amenable to standard stability techniques. The summation in Equation (55), may be written compactly as

$$\sum_{i=1}^{N} \frac{K_i e^{p_i(mT_{sm}-\tau)}}{(1-e^{p_iT_{sm}})r + 1 + e^{p_iT_{sm}}} = \sum_{i=1}^{N} \frac{K_{im}}{(r - r_{im})} \qquad \ldots\ldots\ldots\ldots\ldots \quad 56$$

in which K_{im}, r_{im} are defined by

$$K_{im} = \frac{K_i e^{p_i(mT_{sm}-\tau)}}{1 - e^{p_iT_{sm}}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 57a$$

and

$$r_{im} = \frac{e^{p_iT_{sm}} + 1}{e^{p_iT_{sm}} - 1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 57b.$$

That this summation (Equation (56)) possesses the properties of a partial fraction expansion of a stable rational (though not necessarily minimum phase) function is established by the following argument:

---

[12] The symbol, r, used here is not related to r(t), the symbol for the driving function.

[13] Deleted

Since G(s) is a rational stable function (by assumption) its poles appear as negative real numbers and/or in complex conjugate pairs with negative real parts. Moreover, if the pole, p_i, is real then the corresponding residue, K_i, is real and if the poles, p_i and p_j, appear as a complex conjugate pair then the corresponding residues, K_i and K_j do also. From this, it follows that the poles and residues, r_{im} and K_{im}, conform to the same pattern and, indeed, in exact correspondence with those of G(s). Specifically, if p_i and K_i are real then r_{im} and K_{im} are real also. Similarly, if p_i, p_j and K_i, K_j are corresponding complex conjugate pairs then so are r_{im}, r_{jm} and K_{im}, K_{jm}.

Furthermore, from Equation (57b), each pole, r_{im}, of the summation has a real part given by

$$\mathrm{Re}\left\{r_{im}\right\} = \mathrm{Re}\left\{\frac{e^{p_i T_{sm}} + 1}{e^{p_i T_{sm}} - 1}\right\} \qquad 58$$

which reduces to

$$\mathrm{Re}\left\{r_{im}\right\} = \frac{e^{2T_{sm}[\mathrm{Re}(p_i)]} - 1}{\left|e^{p_i T_{sm}} - 1\right|^2} \qquad 59$$

after a few manipulations. Since $\mathrm{Re}(p_i)$ is negative for all i (because G(s) is stable) it follows from Equation (59) that $\mathrm{Re}(r_{im})$ is also. Therefore, the summation (Equation (56)), is the partial fraction expansion of a stable rational function. Hence it can be expressed as the ratio of two polynomials, that is, by

$$\sum_{i=1}^{N} \frac{K_{im}}{r - r_{im}} = \frac{N_m(r)}{D_m(r)} \qquad 60$$

in which the order of $D_m(r)$ exceeds that of $N_m(r)$. Insertion of this expression for the summation in Equation (55), gives

$$L_m\left(\frac{r+1}{r-1}\right) = \frac{(r-1)^m}{(r+1)^{m-1}} \frac{N_m(r)}{D_m(r)} \qquad \text{61.}$$

On the other hand, the successive application of the z and r-transformations to Equation (42), produces

$$H_m\left(\frac{r+1}{r-1}\right) = 1 + A_m L_m\left(\frac{r+1}{r-1}\right) \qquad 62$$

which on substitution for $L_m\left(\frac{r+1}{r-1}\right)$ from Equation (61) yields

$$H_m\left(\frac{r+1}{r-1}\right) = 1 + A_m \frac{(r-1)^m}{(r+1)^{m-1}} \frac{N_m(r)}{D_m(r)} \qquad 63$$

from which the stability equivalent system is constructed in Figure 12.

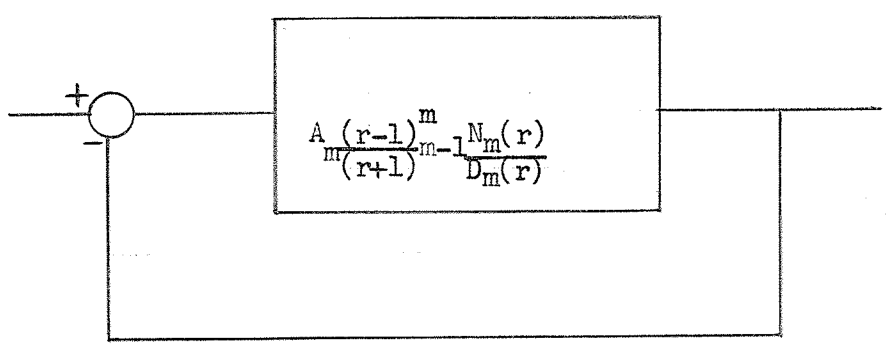


FIGURE 12

THE STABILITY EQUIVALENT SYSTEM

The m^{th} limit cycle is stable if and only if the stability equivalent system (Figure 12) is stable.

The mappings required to obtain this stability equivalent system are conveniently summarized in Figure 13.

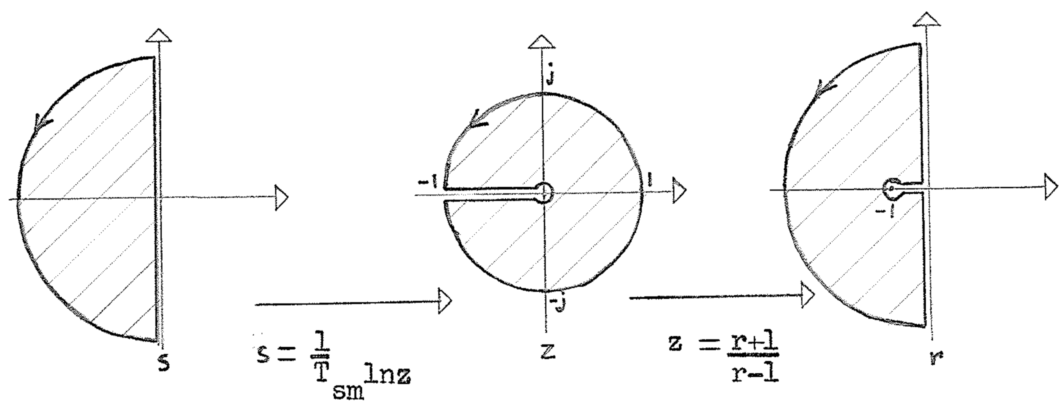


FIGURE 13

TRANSFORMATION APPLIED TO THE VARIATIONAL SYSTEM

(FIGURE 11)

5. ANALYSIS OF THE STABILITY EQUIVALENT SYSTEM

Since only a simple yes or no suffices to answer the question of limit cycle stability, the Routh Test is, perhaps, the most appropriate of the standard analysis techniques.

The Routh Test

From Equation (63), the ordinary Routhian Polynomial, $P_m(r)$, is given by

$$P_m(r) = (r+1)^{m-1} D_m(r) + A_m (r-1)^m N_m(r) \quad \dots\dots\dots\dots 64$$

which when evaluated at the origin gives

$$P_m(0) = D_m(0) + A_m(-1)^m N_m(0) \quad \dots\dots\dots\dots\dots\dots 65.$$

However, Equation (38), gives A_m as

$$A_m = \frac{2M}{\dot{x}_m(-\tau)}$$

or equivalently by

$$A_m = \frac{-2M}{\dot{c}_m(-\tau)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 66$$

which, from Equations (22b and 23b), evaluated at the limit cycle period, T_m, becomes

$$A_m = \frac{-2M}{2M(-1)^m \sum_{i=1}^{N} \frac{K_i e^{p_i(\frac{mT_m}{2}-\tau)}}{1 + e^{p_i \frac{T_m}{2}}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 67.$$

Substitution of T_{sm} for $\frac{Tm}{2}$ in Equation (67) gives

$$A_m = \frac{-1}{(-1)^m \sum_{i=1}^{N} \frac{K_i e^{p_i(mT_{sm}-\tau)}}{1 + e^{p_i T_{sm}}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 68$$

after simplification.

On the other hand, when Equation (55) and Equation (61) are evaluated at the origin and the common term eliminated one obtains

$$(-1)^m \sum_{i=1}^{N} \frac{K_i e^{p_i(mT_{sm}-\tau)}}{1 + e^{p_i T_{sm}}} = (-1)^m \frac{N_m(0)}{D_m(0)} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 69$$

from which

$$N_m(0) = D_m(0) \sum_{i=1}^{N} \frac{K_i e^{p_i(mT_{sm}-\tau)}}{1 + e^{p_i T_{sm}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 70$$

is produced. Substitution of A_m from Equation (68) and $N_m(0)$ from Equation (70) in Equation (65), yields

$$P_m(0) \doteq 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 71.$$

Therefore, the reduced polynomial, $P'_m(r)$, defined by

$$P'_m(r) = \frac{(r+1)^{m-1} D_m(r) + A_m(r-1)^m N_m(r)}{r} \quad \ldots\ldots\ldots \quad 72$$

is the appropriate polynomial to investigate for right-half plane zeroes by the Routh Test.

Though the Routh Test is simple and adequate it suffers from two major disadvantages. First, because of the time-lag there will always be a nested infinity of limit cycles occurring at small periods. Besides the obvious impracti~~b~~cability of testing an infinity of possibilities, each successive one (as T_{sm} tends to zero) has a larger m, so that the expansion of the binomials and the subsequent collection of terms required to manipulate Equation (72) into the proper Routhian form (which displays the coefficients explicitly) becomes quite tedious. Second, the Routh Test gives no indication of the degree of stability of the limit cycle.

<u>Stability of the limiting case</u>. The first aforementioned difficulty with the Routh Test is somewhat alleviated by the following discussion of the limiting form (as m increases) of the

denominator of the closed loop transfer function of the stability equivalent system (Figure 12).

Substitution of Equation (68) and Equation (55) in Equation (62) gives this function as

$$H_m\left(\frac{r+1}{r-1}\right) = 1 + \frac{(-1)^m (r-1)^m \sum_{i=1}^{N} \dfrac{K_i e^{p_i(mT_{sm} - \tau)}}{(1-e^{p_iT_{sm}})r + 1 + e^{p_iT_{sm}}}}{(r+1)^{m-1} \sum_{i=1}^{N} \dfrac{K_i e^{p_i(m\,T_{sm} - \tau)}}{1 + e^{p_iT_{sm}}}} \quad \ldots\ldots \quad 73$$

However, it will be more convenient to work with $H'_m\left(\frac{r+1}{r-1}\right)$ given by

$$H'_m\left(\frac{r+1}{r-1}\right) = (r+1)^{m-1} \sum_{i=1}^{N} \frac{K_i e^{p_i(mT_{sm} - \tau)}}{1 + e^{p_i T_{sm}}} +$$

$$(-1)^{m+1} (r-1)^m \sum_{i=1}^{N} \frac{K_i e^{p_i(mT_{sm} - \tau)}}{(1 - e^{p_iT_{sm}})r + 1 + e^{p_iT_{sm}}} \quad \ldots \quad 74$$

a function which has the same zeros as $H\left(\frac{r+1}{r-1}\right)$. Introduce d_m, defined by

$$d_m = \frac{m\,T_{sm} - \tau}{T_{sm}} \quad \ldots\ldots\ldots\ldots \quad 75$$

which from Figure 9 satisfies

$$0 \leqslant d_m < 1 \quad \ldots\ldots\ldots\ldots \quad 76$$

for all m. From the above two equations it is evident that as m increases T_{sm} tends to zero, such that mT_{sm} remains finite.

Insertion of Equation (75) in Equation (74) gives

$$H'_m\left(\frac{r+1}{r-1}\right) = (r+1)^{m-1}\sum_{i=1}^{N}\frac{K_i e^{p_i T_{sm} d_m}}{1+e^{p_i T_{sm}}} +$$

$$(-1)^{m+1}\ (r-1)^m \sum_{i=1}^{N}\frac{K_i e^{p_i T_{sm} d_m}}{(1-e^{p_i T_{sm}})r+1+e^{p_i T_{sm}}} \qquad \ldots\ldots\ 77$$

which reduces to

$$H'_m\left(\frac{r+1}{r-1}\right) = (r+1)^{m-1}\sum_{i=1}^{N}\frac{K_i}{2} + (-1)^{m+1}(r-1)^m\sum_{i=1}^{N}\frac{K_i}{2-p_i T_{sm} r} \qquad \ldots\ldots\ 78$$

for large m (small T_{sm}). If G(s) falls off more quickly than $\frac{1}{s}$ for large s then $\sum_{i=1}^{N} K_i$ vanishes as shown by the following

<u>Lemma</u>: If G(s) asymptotically approaches $\frac{K}{s^n}$ for large s, where K is a constant and n is a positive integer greater than unity, then $\sum_{i=1}^{N} K_i p_i$ vanishes for k equal to 1, 2, 3...n-1.

<u>Proof</u>: From Equation (17) the step response of G(s), q(t), is given by

$$q(t) = -\sum_{i=1}^{N}\frac{K_i}{p_i} + \sum_{i=1}^{N}\frac{K_i e^{p_i t}}{p_i}$$

from which the k^{th} derivative is given by

$$\frac{d^k}{dt^k}q(t) = \sum_{i-1}^{N} K_i p_i^{k-1} e^{p_i t} \qquad \ldots\ldots\ 79$$

which when evaluated at the origin produces

$$\frac{d^k}{dt^k}q(0) = \sum_{i=1}^{N} K_i p_i^{k-1} \qquad \ldots\ldots\ 80.$$

On the other hand, the initial value theorem yields

$$\frac{d^k}{dt^k} q(0) = \lim_{s\to\infty} s \left\{ s^k \frac{G(s)}{s} \right\} \qquad \text{......... 81}$$

provided all the lower order derivatives are zero at the origin (zero initial conditions). Because of the asymptotic behaviour of G(s) Equation (81) reduces to

$$\frac{d^k}{dt^k} q(0) = K \lim_{s\to\infty} \frac{1}{s^{n-k}} \qquad \text{......... 82}$$

and in particular

$$\frac{d^k}{dt^k} q(0) = 0 \qquad \text{......... 83}$$

for all k less than n. When the right-hand side of Equation (80) is set equal to zero in accordance with Equation (83),

$$\sum_{i=1}^{N} K_i p_i^{k-1} = 0 \qquad \text{......... 84}$$

obtains for all k less than n. This is the desired result.

Thus if $G(s)$ falls off more quickly than $\frac{1}{s}$ for large s then $\sum_{i=1}^{N} K_i = 0$ and Equation (78) reduces to

$$H_m'\left(\frac{r+1}{r-1}\right) = (-1)^{m+1} (r-1)^m \frac{\sum_{j=1}^{N} K_i \prod_{i\neq j}^{N} (2 - p_i T_{sm} r)}{\prod_{i=1}^{N} (2 - p_i T_{sm} r)} \qquad \text{......... 85}$$

in which the summation is rewritten as the ratio of two polynomials. The coefficients of the highest and second highest powers in r in the Routhian Polynomial (which is the numerator of Equation (85) are of opposite sign for sufficiently large m; therefore the associated limit cycles (for large m) are unstable. This result is established for the case in which all p_i are negative real as follows:

The highest power of r in the numerator of Equation (85) is r^{m+N-1}. The coefficient of this term, A, is given by

$$A = T_{sm}^{N-1} \sum_{j=1}^{N} K_j \prod_{i \neq j}^{N} (-p_i) \qquad \text{.........} \quad 86$$

and the coefficient of r^{m+N-2}, B, is given by

$$B = 2T_{sm}^{N-2} \sum_{j=1}^{N} K_j \left[\prod_{i \neq j,2}^{N} (-p_i) + \prod_{i \neq j,3}^{N} (-p_i) + \ldots \prod_{i \neq j,N}^{N-1} (-p_i) \right] - mT_{sm}^{N-1} \sum_{j=1}^{N} K_j \prod_{i \neq j}^{N} (-p_i) \qquad \text{..........} \quad 87$$

which for sufficiently large m reduces to

$$B = -m\, T_{sm}^{N-1} \sum_{j=1}^{N} K_j \prod_{i \neq j}^{N} (-p_i) \qquad \text{..........} \quad 88.$$

Since A and B given by Equation (86) and Equation (88) respectively are opposite in sign the assertion is established. Hence for all m larger than some critical value all limit cycles are unstable. The main weakness in this result is that the critical m is poorly delineated.

It is expected that a similar result may be obtained for the case in which complex p_i are allowed although no proof is attempted here.

Equations (89) to (94) inclusive and page 44 have been deleted.

after a few manipulations are performed. Again, the first two coefficients are of opposite sign (see Inequality (81b)), hence every limit cycle with period smaller than T_c is unstable. The weakness in this result is that the critical period, T_c, is poorly delimited. Nevertheless, a necessary condition that the Taylor Expansion be valid requires that

$$|p_i T_{sm}| \ll 1$$

be satisfied for all i and m. This becomes

$$\frac{T_c}{2} \ll \frac{1}{|p_i|} \qquad \text{.................................} \qquad 94$$

for all i in terms of the critical half period. From this, a rough idea of T_c may be obtained.

<u>Comment</u>: The existence of a critical period has been established for all systems (under consideration here) which behave asymptotically as $\frac{1}{s}$ or $\frac{1}{s^2}$, that is, for n equal to one or two. It may be possible (perhaps by mathematical induction) to establish this for all n, though such a proof is not attempted here.

<u>The Nyquist Methods</u>

Either the Nyquist or the dual Nyquist[13] Method when applied to the open loop function, $A_m L_m \left(\frac{r+1}{r-1}\right)$, is free from the manipulative difficulties inherent in the Routh Test. On the other hand, each Nyquist Diagram intersects its respective critical point because

---

[13] The factor, $\left(\frac{r-1}{r+1}\right)^{m-1}$, along the Bromwich Contour behaves exactly the same as a time-delay factor so that the dual Nyquist Method does apply. For a discussion of the dual Nyquist Diagram see:

P. Jones, <u>Stability of Feedback System Using Dual Nyquist Diagrams</u>, I.R.E. Transactions, vol. CT-1, #1 p. 35, 1954.

M. Satche, Journal of Applied Mechanics, December, 1949, pp. 419-20.

the Routhian Polynomial has a zero at the origin. This means that no estimate of the degree of stability of the limit cycle can be obtained from either the Nyquist or the dual Nyquist Method. However, such an estimate is provided by the Nyquist Diagram for the reduced function, $H'_m(r)$, defined by

$$H'_m(r) = \frac{1 + A_m L_m \frac{(r+1)}{(r-1)}}{r} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ 95.$$

Unfortunately, the dual Nyquist Diagram for $H'_m(r)$ is complicated by the $\frac{1}{r}$ factor and the method loses much of its efficacy.

Caveat: It must be emphasized that whatever method is employed, each limit cycle is tested individually, not collectively. Thus, for each limit cycle, either a new Routhian Polynomial must be computed or a new Nyquist Diagram constructed.

Though both the Routh and Nyquist Methods are intrinsically cumbersome, it is possible, with any of the Nyquist Methods for a definite trend to emerge after constructing only a few diagrams. Because of this, the Nyquist Methods appeared to be superior to the Routh Test from a system point of view.

6. AN EXAMPLE

The example that initiated the study (Figure 1) is now analyzed by the methods developed in this chapter. The values of the system constants are given by

$$M = 1, \ \tau = 1, \ p_1 = -1, \ K_1 = -1, \ p_2 = 0, \ K_2 = 1$$

which when inserted in Equations (25a, b) produce:

$$\mathrm{Re}\left\{H_m\left(\frac{T}{2}\right)\right\} = (-1)^m\left(-2 + \frac{2e^{1-m\frac{T}{2}}}{1+e^{-\frac{T}{2}}} + (2m-1)\frac{T}{4}\right) \quad \text{96a}$$

and

$$\mathrm{Im}\left\{H_m\left(\frac{T}{2}\right)\right\} = 2(-1)^m\left(1 - \frac{e^{1-m\frac{T}{2}}}{1+e^{-\frac{T}{2}}}\right) \quad \text{96b}$$

for the m^{th} Hamel Function on the range,

$$\frac{1}{m} \leqslant \frac{T}{2} < \frac{1}{m-1} \ .$$

The first four Hamel Functions are plotted in tandem in Figure 13.

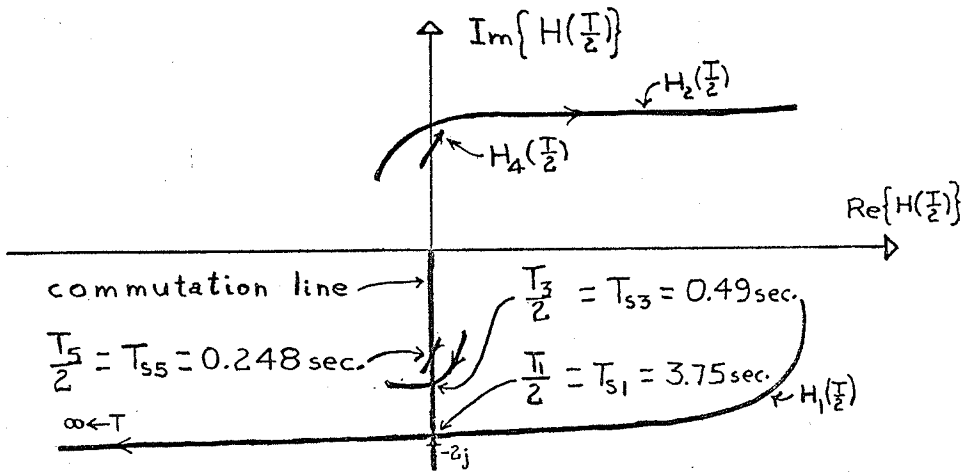


FIGURE 13

HAMEL LOCUS FOR EXAMPLE

The form of the Hamel Locus strongly suggests that this system has infinitely many limit cycles. This refutes the claim of Gille et al. that such a system has only one limit cycle.[14] It is true, however, that this system has only one stable and therefore one physical limit cycle. Indeed, that the limit cycle with half period of 3.75 seconds is stable is readily established by the Routh Test as follows:

From Equation (60) and Equation (56) it is evident that $\frac{N_m(r)}{D_m(r)}$ may be expressed by

$$\frac{N_m(r)}{D_m(r)} = \sum_{i=1}^{N} \frac{K_i e^{p_i(mT_{sm}-\tau)}}{(1 - e^{p_i T_{sm}})r + 1 + e^{p_i T_{sm}}} \quad \ldots\ldots\ldots\ldots\ldots \; 97$$

which when evaluated at the first limit cycle (half period 3.75 seconds) produces

$$\frac{N_1(r)}{D_1(r)} = \frac{0.5(r + 0.917)}{r + 1.048} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 98$$

after simplification. Substitution of this expression in Equation (72), produces

$$P_1'(r) = \frac{r + 1.048 + A_1(r - 1)(0.5)(r + 0.917)}{r} \quad \ldots\ldots\ldots\ldots \; 99,$$

the reduced Routhian Polynomial for the first limit cycle. The gain constant, A_1, is readily computed by evaluating the numerator of the above expression at the origin and by setting the result equal to zero.[15] When this equation is solved for A_1, the result is

---

[14] Gille et al., op. cit., pp 440-1.

[15] For a proof that the numerator of the right-hand side of Equation (99) vanishes at the origin see pp.37-9.

$$A_1 = \frac{(-1)^2(1.048)}{(0.5)(0.917)} = 2.28 \qquad \ldots\ldots\ldots\ldots \quad 100$$

which when reinserted in Equation (99) gives

$$P_1' = 1.14r + 0.905 \qquad \ldots\ldots\ldots\ldots \quad 101$$

after simplification. Since this reduced Routhian Polynomial is obviously stable, the first limit cycle is also stable.

When a similar analysis is applied to the adjacent limit cycle (half period 0.49 seconds) the reduced Routhian Polynomial simplifies to

$$P_3'(r) = 4.43r^3 - 8.14r^2 + 6.96r + 17.38 \qquad \ldots\ldots\ldots\ldots \quad 102$$

which is obviously unstable; therefore, the associated limit cycle is unstable.

Furthermore, $G(s)$ satisfies the conditions set forth in "Stability of the limiting case"; hence for sufficiently large m all limit cycles are unstable. In deducing this result the approximations

$$1 + e^{p_i T_{sm}} \doteq 2 \qquad \ldots\ldots\ldots\ldots \quad 103a$$

and

$$1 - e^{p_i T_{sm}} \doteq p_i T_{sm} \qquad \ldots\ldots\ldots\ldots \quad 103b$$

were assumed to hold for all i. Since p_2 is zero these equations are satisfied exactly for i equal to 2. For i equal to 1 the left-hand sides are

$$1 + e^{p_1 T_{s3}} = 1 + e^{(-1)(0.248)} = 1.78 \quad \text{and} \quad 1 - e^{p_1 T_{s3}} = 1 - e^{(-1)(0.248)} = 0.22$$

assuming a critical m of 3. If 10% error may be tolerated in Equation (103a,b) then all limit cycles with m greater than 1 are unstable. This conclusion is in agreement with the experimental facts.

An interesting sidelight is that this system may be represented in a "phase plane" in which the limit cycles do not alternate in stability in accord with Poincare's Rule. This apparent contradiction, however, is easily explained: The presence of time-lag means that the state space has infinitely many state variables. Consequently, the "phase plane" representation is, in fact, only a phase sub-plane of the entire system. Poincare's Rule, however, applies only to true state or phase planes.[16]

7. THE FAILURE OF LOEB'S RULE

In the derivation of his criterion, Loeb effectively assumed that the solution to the variational equation was a damped sinusoid at the limit cycle frequency.[17] It will be established in this section, however, that this assumption is invalid for each of the limit cycles in the previous example, except the lowest frequency one. This is why the rule works for the lowest frequency limit cycle, yet, fails for all others.

The above statements may be deduced from the transient response of the associated variational system (Figure 11). However, the autonomous solution for $\acute{u}_m(t)$ presents some difficulty as the z-transform method for finding the time response of the output cannot be applied directly in the autonomous case.

---

[16] This apparent contradiction together with its explanation was first observed by Professor R. A. Johnson.

[17] J. Loeb, op. cit., 1951.

On the other hand, the z-transform is applicable to this system with zero initial conditions,[18] and with the non-zero driving function, a unit delta function at the time origin. Furthermore, this driven system is autonomous except at the time origin. Therefore, for all positive time, the characteristic modes of $u^*(t)$ of the driven system are the same as those of the autonomous system. Indeed, the following solution to this driven system is sufficient to establish the invalidity of the Loeb assumption.

Since the z-transform of the unit delta function is unity, the z-transform of the output, $c_m(t)$, for the m^{th} variational system is given by

$$C_m(z) = \frac{A_m L_m(z)}{1 + A_m L_m(z)} \quad \text{......} \quad 104$$

in which A_m and $L_m(z)$ are given by Equation (68), and Equation (53), respectively. Substitution of the appropriate values in Equation (104) for the first limit cycle in the previous example gives

$$C_1(z) = \frac{(2.29)(0.936z + 0.0404)}{(z+1)(z+0.116)} \quad \text{......} \quad 105a$$

after simplification or

$$C_1(z) = \frac{2.32}{z+1} - \frac{0.177}{z+0.116} \quad \text{......} \quad 105b$$

in a partial fraction expansion. The output at a sampling instant is determined from[19]

---

[19] Kuo, op. cit., pp. 66-8.

[18] For the sake of simplicity it has been assumed that there is no memory storage in the delay unit from the time origin to τ seconds later.

$$c_m(nT_{sm}) = \frac{1}{2\pi j}\oint_\Gamma C_m(z)z^{n-1}dz \quad \ldots\ldots\ldots\ldots \quad 106$$

where Γ encloses all the singularities of the integrand and n is a non-negative integer. The evaluation of this inversion integral yields

$$c_1(0) = 0 \quad \ldots\ldots\ldots\ldots \quad 107a$$

for n equal to zero and

$$c_1(3.75n) = 2.32(-1)^{n-1} - (0.177)(-0.116)^{n-1} \quad \ldots\ldots\ldots\ldots \quad 107b$$

for positive n. From a physical standpoint, Equation (107a) is obvious because the time-delay in the forward path delays the effect of the delta function at the output by τ seconds.

Since the system is autonomous for positive time, the error signal is the negative of Equation (107b) which is

$$u_1(3.75_n) = -2.32(-1)^{n-1} + (0.177)(-0.116)^{n-1} \ldots\ldots\ldots \quad 108$$

for positive n. This is the solution to the variational equation at the sampling instants for a particular set of initial conditions. By itself, Equation (108) does not specify a unique time function for the unsampled error signal, $u_1(t)$. However, either the modified z-transform or submultiple sampling may be used to approximate $u_1(t)$ to any desired degree of precision, though such refinement will not be required for what follows.

From a lemma in Chapter II it is known that $K\dot{x}_1$ (K is a constant) is always a solution to the variational system. By identification, $K\dot{x}_1$ is represented by the first term of Equation (108). Loeb, in his analysis, completely ignored this steady state part of the solution and, although this was conceptually incorrect, it did not lead to false conclusions because the stability of the variational system (and hence the limit cycle) does not depend on this steady state solution.

To approximate the unsampled error signal, $u_1(t)$, from the sampled error signal, $u_1(3.75n)$, the first term in Equation (108), is replaced by

$$\frac{(2.32)\dot{x}_1(t)}{\dot{x}_1(3.75)}$$

with zero error since K was chosen so that $K\dot{x}_1$ was equal to the first term of Equation (108) at the first sampling instant. The second term is aproximated by a damped sinusoid at the limit cycle frequency of $\frac{\pi}{3.75}$ or 0.839 rads./sec. This procedure gives

$$u_1(t) \doteq \frac{(2.32)\dot{x}_1(t)}{\dot{x}_1(3.75)} - 0.177e^{2.15}e^{-0.575t}\cos(0.839t) \;\ldots\ldots\ldots\; 109$$

which is identical to Equation (108) at the sampling instants.[20] Because of this it follows that Equation (109) is at least a rough approximation to Equation (108). Thus, Loeb's assumption justified for the first limit cycle.

Insertion of the appropriate constants in Equation (104), for the next limit cycle (half period 0.49 seconds) yields

$$C_3(z) = \frac{8.9(0.375z + 0.012)}{(z+1)(z+0.0309)(z^2 - 2.645z + 3.32)} \;\ldots\ldots\ldots\text{110a}$$

which when expressed as a partial fraction expansion becomes

$$C_3(z) = \frac{0.479}{z+1} + \frac{0.009}{z+0.0309} - j\frac{0.508e^{-j28.3^\circ}}{z - 1.82e^{-j43.5^\circ}} + j\frac{0.508e^{j28.3^\circ}}{z - 1.82e^{j43.5^\circ}}$$

$$\ldots\ldots\ldots\; \text{110b}$$

---

[20] To prove this, replace t by 3.75n in the second term of Equation (109). Since the first term is exact the result follows immediately.

which on inversion by the negative of Equation (106) yields

$$u_3(0.49n) = -0.479(-1)^{n-1} + 0.009(-0.0309)^{n-1} + 1.016(1.82)^{n-1}\sin(0.759n - 71.8^\circ) \quad \ldots\ldots \quad 111$$

the first term of which represents the steady state solution as before. However, the remaining terms cannot be approximated by a single damped sinusoid at the limit cycle frequency because the last and dominant term (the amplitude of which grows with time) does not change sign at the sampling frequency.

Unlike the intuitive discussion justifying Loeb's assumption for the first limit cycle, the above argument constitutes a proof that Loeb's assumption is false for the second limit cycle. Furthermore, it is reasonable to assume (though no proof is attempted) that the rule fails for the same reason for the remaining limit cycles.

However, if the intuitive analysis used for the first limit cycle is pursued for this second limit cycle, further insight is gained into the failure of Loeb's Rule. Thus, the continuous error signal, $u_3(t)$, is now approximated by

$$u_3(t) \doteq \frac{-(0.479)\dot{x}_3(t)}{\dot{x}_3(0.49)} - 0.009e^{3.48}e^{-7.11t}\cos(6.41t) + 1.016e^{-0.6}e^{1.22t}\sin(1.55t - 71.8^\circ) \quad \ldots\ldots\ldots\ldots \quad 112$$

which (as before) is identical to Equation (111) at the sampling instants. It is to be noted that this solution differs materially from the previous one in that the last term is not at the limit cycle frequency of 6.41 rads/second. This fact explains why the rule fails. If only the first and second terms were present in Equation (112), then the limit cycle would be stable as predicted by the rule. The rule, however, takes no account of the third term which is not at the limit cycle frequency but which actually makes the limit cycle unstable.

It is significant that the failure of Loeb's Rule in this example is not due to any error in the describing function approximation. Indeed, a sample calculation by the describing function method gave 7.47 seconds for the period of the first limit cycle (worst case) as compared to 7.50 seconds from the Hamel Locus. Since this is an error of less than 0.5%, it is obvious that the describing function approximation is valid for this example.

8. LOEB'S RULE AND THE TSYPKIN LOCUS RULE

In the previous example, the fact that the first limit cycle is stable and the others unstable does not contradict the following rule derived by Gille et al. for the Tsypkin Locus.[21] It states that an intersection from above the commutation line, for increasing frequency, always represents an unstable limit cycle whereas one from below may represent either a stable or an unstable limit cycle; further investigation is always required to decide this latter case.

---

[21] Gille et al., op. cit., pp. 481-83.

Therefore, in all relay systems for which the describing function method closely approximates the Tsypkin analysis, a positive vector product (which corresponds to an intersection from below by the Tsypkin Locus) does not guarantee a stable limit cycle as asserted by Loeb. Indeed, in the absence of further information, a positive vector product gives no stability information whatsoever. This certainly agrees with the results of the previous example in which every limit cycle has a positive vector product associated with it; yet, one limit cycle is stable and the others are unstable. This rule for the Tsypkin Locus thus exposes a major failing in Loeb's Criterion in relay systems and thereby casts the shadow of doubt upon the positive vector product in all systems.

Unfortunately, Gille et al. employed specious reasoning to derive this rule. For example, in the symbolism of this thesis, it is suggested therein that the zero at s equal to jw_m, where w_m is the limit cycle frequency, be removed from $1 + A_m L^*_m(jw)$ to create the new function, $S_m(jw)$, given by

$$S_m(jw) = \frac{1 + A_m L^*_m(jw)}{s - jw_m} \qquad \text{.............................} \; 113$$

which is to be examined for right-half plane zeroes by the Nyquist Test. This latter function, namely $S_m(jw)$, is not a sampled function because it is not periodic in its argument. Thus, in applying the Nyquist Test, the entire right-half plane must be encompassed, not simply the primary strip of the right-half plane (which is sufficient for a sampled function). But if the entire imaginary axis must be traversed, it is an empty strategy to remove the single zero at s equal to jw_m since $1 + A_m L_m^*(s)$, by virtue of its periodicity, has infinitely many zeroes on the imaginary axis positioned at s equal to $\pm jw_m$, $\pm j3w_m$, $\pm j5w_m \cdots$. Hence, the derivation is invalid. This does not mean that the rule is invalid; in fact, the following argument lends considerable support to the rule.

Perhaps, the simplest way to attack this problem is to apply the Nyquist Test to the sampled function, $1 + A_m L_m^*(s)$, for which the appropriate s-plane contour (denoted by C) is shown in Figure 14. The detours at $\pm jw_m$

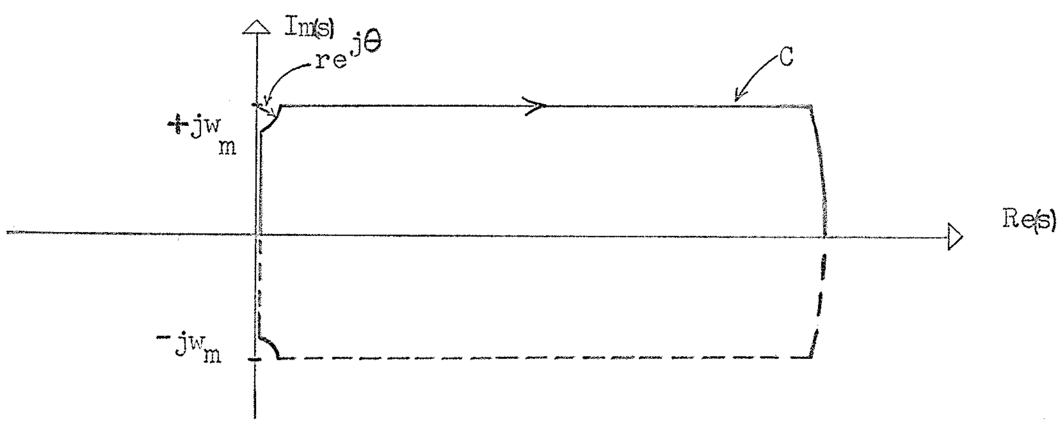


FIGURE 14

THE s-PLANE CONTOUR FOR $1 + A_m L_m^*(s)$

are always needed to prevent the contour from passing through the zeroes of $1 + A_m L_m^*(s)$ situated at these points.[21] Additional detours are required

[21] In the Nyquist Test, the s-plane contour must not pass through any poles or zeroes of the function being tested.

if L(s) has any poles on the imaginary axis. The limit cycle will be stable, if as s traces out C, $L_m^*(s)$ does not encircle the critical point, $-\frac{1}{A_m}$; it will be unstable otherwise. Some general properties of the sampled locus are now deduced.

In the neighborhood of the critical point, the sampled locus may be approximated by the two term Taylor Expansion at s equal to jw_m, given by

$$L_m^*(jw_m+re^{j\theta}) \doteq L_m^*(jw_m) + re^{j\theta}\frac{dL_m^*}{ds}(jw_m) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 114$$

in which $re^{j\theta}$ is defined in Figure 14. Straightforward differentiation and substitution in Equation (23a) and Equation (50) establishes

$$\frac{dL_m^*}{ds}(jw_m) = \frac{w_m}{2M}\frac{d}{dw}\left(\mathrm{Re}\left\{H_m\left(\frac{T_m}{2}\right)\right\}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 115$$

which, since $\frac{w_m}{2M}$ is a positive real number, proves that $\frac{dL^*}{ds}(jw_m)$ is a real number whose sign is the same as $\frac{d}{dw}(\mathrm{Re}H_m(\frac{T_m}{2}))$. The sign of the latter derivative is easily read from the Hamel Locus. Furthermore, the real part of the Hamel Locus is equal to the imaginary part of the Tsypkin Locus. Therefore, the previous equation may be rewritten in terms of the Tsypkin Locus as

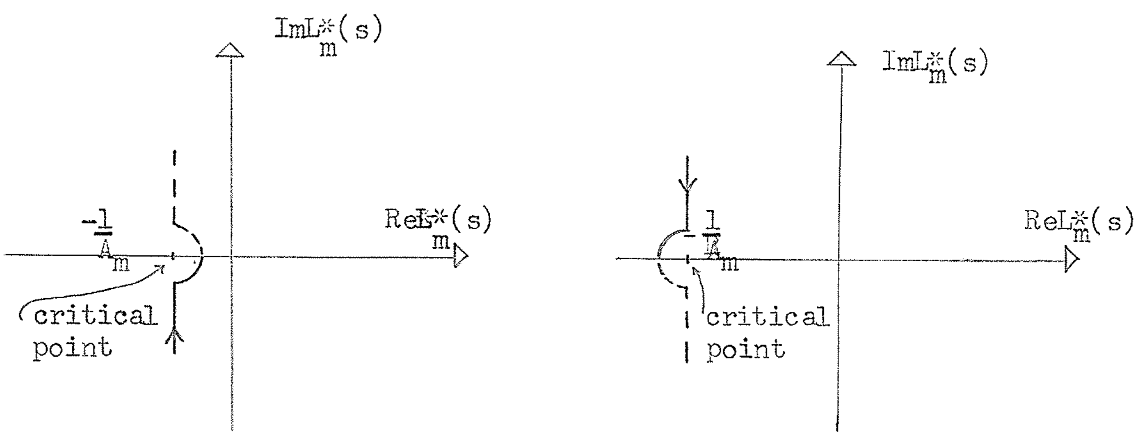
$$\frac{d}{ds}L_m^*(jw_m) = \frac{w_m}{2M}\frac{d}{dw}\left(\mathrm{Im}\left\{\Lambda(w_m)\right\}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 116$$

The sign of $\frac{d}{dw}(\mathrm{Im}\left[\Lambda(w_m)\right])$ is positive if the Tsypkin Locus approaches the commutation line from below for increasing w and negative in the other case, a fact which is apparent for any particular limit cycle by inspection of the Tsypkin Locus.[22] From Equation (114), the behaviour of the sampled locus in the neighborhood of the critical point is sum-

[22] In the Hamel representation, the derivative is positive if the locus cuts the commutation line from the right for increasing T; it is negative in the opposite case.

marized graphically in Figure 15.

Another general fact about the sampled locus (which is easily proved by setting s equal to zero in Equation (50)) is that $L_m^*(0)$ is a real number. This means that the sampled locus intersects the real axis at s equal to zero. Now, if (1) this intersection lies to the right of the critical point and (2) the sampled locus does not cut the real axis for any other values of s as C is traced out then the m^{th} limit cycle is stable for $\frac{dL_m^*}{ds}(jw_m)$ positive (Figure 15a) and unstable for $\frac{dL_m^*}{ds}(jw_m)$ negative (Figure 15b). This is the essence of the rule derived by Gille et al. with two sufficiency assumptions appended.



a) $\frac{dL_m^*}{ds}(jw_m) > 0$

b) $\frac{dL_m^*}{ds}(jw_m) < 0$

Note: The Tsypkin Locus, $\Lambda(w)$, cuts the commutation line from below (not shown).

Note: The Tsypkin Locus, $\Lambda(w)$, cuts the commutation line from above (not shown).

FIGURE 15

BEHAVIOUR OF $L_m^*(s)$ NEAR THE CRITICAL POINT

There is (at least) one general case for which the first assumption is not required. Suppose G(s) (refer to Figure 8) contains a single integration then if p_N is taken as zero (without loss of generality) the corresponding residue, K_N, will be a positive real number. To prove this, observe that since G(s) is assumed to be a rational, stable, and minimum phase function it may be expressed as

$$G(s) = \frac{\sum_{i=0}^{N-2} a_i s^i}{s\sum_{j=0}^{N-2} b_j s^j} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 117$$

in which a_i, b_j are positive real numbers for all i and j. The K_N th residue is given by

$$K_N = \lim_{s\to 0}\left\{sG(s)\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 118$$

which by virtue of Equation (117) becomes

$$K_N = \frac{a_o}{b_o} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 119$$

in the limit. Since a_o and b_o are both positive real numbers, K_N is also. On the other hand, substitution of p_N equal to zero in Equation (50), yields

$$L_m^*(s) = \sum_{i=1}^{N-1} \frac{K_i e^{-p_i\tau} e^{-mT_{sm}(s-p_i)}}{1 - e^{-T_{sm}(s-p_i)}} + \frac{K_N e^{-mT_{sm}s}}{1 - e^{-T_{sm}s}} \quad \ldots\ldots\ldots\ldots \quad 120$$

for the sampled function when G(s) contains a single integration. As s tends to zero in Equation (120), $L_m^*(0)$ tends to $+\infty$ since K_N is a positive real number. Hence, in this case, $L_m^*(0)$ is always to the right of the critical point (which is a negative real number); therefore, the first assumption is not required. It should be noted that this remark applies to the previous example (since it has one integration in G(s)).

As the discussion is not intended to be an exhaustive study of this rule for the Tsypkin Locus, no further proofs will be attempted here. The purpose of this discussion is only to expose certain fallacies in the existing derivation and to lay a firm foundation for further research. To this end, two possible "theorems" are proposed for further study:

> "Theorem 1": If G(s) has one integration then an intersection from above the commutation line for increasing w by the Tsypkin Locus always implies an unstable limit cycle at the frequency of the intersection. (This theorem would eliminate the second assumption for intersections from above.)

> "Theorem 2": If G(s) has one integration then the lowest frequency intersection of the Tsypkin Locus with the commutation line represents a stable or unstable limit cycle accordingly as the intersection occurs from below or above the commutation line for increasing w. (This theorem would eliminate the second assumption for the first or lowest frequency intersection.)

No proofs are offered for these "theorems", nor is it claimed that they are, in fact, true. Nevertheless, they could provide a convenient starting point for further work. In this connection, Gille et al. cite a Russian reference which might be useful, though it was not reviewed by the author of this thesis.[23]

---

[23] G. Nejmark, O periodicheskikh rezhimakh i ustojchivasti relejnykh sistem, Automatika i Telemekhanika, 14(5): 556-569 (1953).

CHAPTER IV

DISCUSSION OF RESULTS

1. SUMMARY

The failure of Loeb's Rule in the simple control system described in Chapter I provided the main justification for this thesis. Chapter II was devoted to a general discussion on limit cycle stability and served as a base for Chapter III wherein an exact analysis was developed for limit cycles and limit cycle stability for autonomous time-lag ideal relay control systems. This analysis was then used to solve the original problem of Chapter I; the resulting solution provided considerable insight into why the rule failed. Last of all, a constructive criticism of a rule for the Tsypkin Locus was offered and some implications of this rule were discussed.

2. CONCLUSIONS

The first conclusion is that Loeb's Rule failed in the example cited in Chapter I because of unjustified assumption in the derivation of the rule. In particular, it was proved that the solution to the variational equation was not always a damped sinusoid at the limit cycle frequency as Loeb had assumed in his derivation. Therefore, neither the slight error in the describing function approximation nor the presence of time lag contributed in any essential way to this failure. The second important result was the development, in Chapter III, of an exact limit cycle analysis for time-lag ideal relay control systems to replace the approximate describing function - Loeb method. Finally, a rule for the Tsypkin Locus suggested that a positive vector product per se in the describing function - Loeb method was an inconclusive stability test for ideal relay control systems and therefore, by implication, for other systems as well.

3. FURTHER STUDY

Some topics for further research are:

1. the extension of the method of Chapter III to other on-off elements and ultimately, with the aid of Mohammed's work, to the general nonlinear element containing saturation, dead zone, and hysteresis,[1]

2. a detailed study into possible alternatives to Loeb's Rule for nonlinearities other than those in (1) above, using the results of Chapter II as a starting point,

3. an investigation into the behaviour of the limiting form of the Routhian Polynomial for the stability equivalent

[1] A.Mohammed, op. cit..

system for systems in which G(s) falls off more quickly than $\frac{1}{s^2}$ (see Chapter III) and

4. a complete examination of the rule for the Tsypkin Locus starting with the "theorems" indicated in Chapter III.

BIBLIOGRAPHY

Coddington, E. A. and N. Levinson. Theory of Ordinary Differential Equations. New York: McGraw-Hill Book Company, Inc., 1955.

Cosgriff, R. L. "Application of Linear Differential Equations with Periodic Coefficients in the Study of Non-linear Phenomena" Proc. of First International Congress of the I.F.A.C. Vol. 2, 1960, pp. 883-87.

Gibson, John E. Nonlinear Automatic Control. New York: McGraw-Hill Book Company, Inc., 1963.

Gille, J. C., M. J. Pélegrin, and P. Decaulne. Feedback Control Systems. New York: McGraw-Hill Book Company, Inc., 1959.

Grensted, P. E. W. "The Frequency Response Analysis of Non-Linear Systems," Proc. I.E.E., Vol. 102, part C, 1955, pp. 244-255.

Grensted, P. E. W. "Analysis of the Transient Response of Non-Linear Control Systems," A.S.M.E. Trans. 80, 1958, pp. 427-32.

Grensted, P. E. W. "Frequency Response Methods Applied to Non-Linear Systems" In Progress in Control Engineering, Vol. 1, pp. 105-139, Ed. R. H. Macmillan et al..New York: Academic Press Inc., 1962.

Hayashi, Chihiro. Nonlinear Oscillations in Physical Systems. New York: McGraw-Hill Book Company, Inc., 1964.

Johnson, R. A. "State Space and Systems Incorporating Delay," Electronics Letters, Vol. 2, No. 7, July 1966, pp. 277-78.

Jones, Paul. "Stability of Feedback Systems Using Dual Nyquist Diagrams," Trans. I.R.E., Vol. CT-1, No. 1, 1954, p. 35.

Kuo, Benjamin C. Analysis and Synthesis of Sampled-Data Control Systems. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.

Loeb, J. "Phénomènes Héréditaire dans les Servomécanismes; un Critérium Général de Stabilité," Annales des Télécommuncations, 6(12): 346-356, 1951.

Loeb, J. "Recent Advances in Nonlinear Servo Theory." (1953) In Frequency Response ed. R. Oldenburger, New York: The Macmillan Book Company, 1956, pp. 260-67.

Minorsky, N. Nonlinear Oscillations. Princeton, N.J.: D. Van Nostrand Company, Inc., 1962.

Mohammed, Auyuab. "Steady-State Oscillations and Stability of On-Off Feedback Systems," Ph. D. Thesis, University of British Columbia, Department of Electrical Engineering, April 1965.

Mufti, I. H. "A Method For the Exact Determination of Periodic Motions in Relay Control Systems," Proc. of J.A.C.C. Troy, New York, June 1965.

Nejmark, G. "O periodicheskikh rezhimakh i ustojchivosti relejnykh sistem," Avtomatika i Telemekhanika, 14(5): 556-569 (1953).

Pontryagin, L. S. Ordinary Differential Equations. Reading Mass.; Addison-Wesley Publishing Company, Inc., 1962.

Satche, M. Journal of Applied Mechanics, Dec. 1949, pp. 419-20.

Spiegel, Murray R. Theory and Problems of Complex Variables. New York: Schaum Publishing Company, 1964.

APPENDIX

Theorem: The variational equation obtained by the operational approach is equivalent to those obtained by the state approach for a nonlinear feedback control system without time-lag (Figure 16).

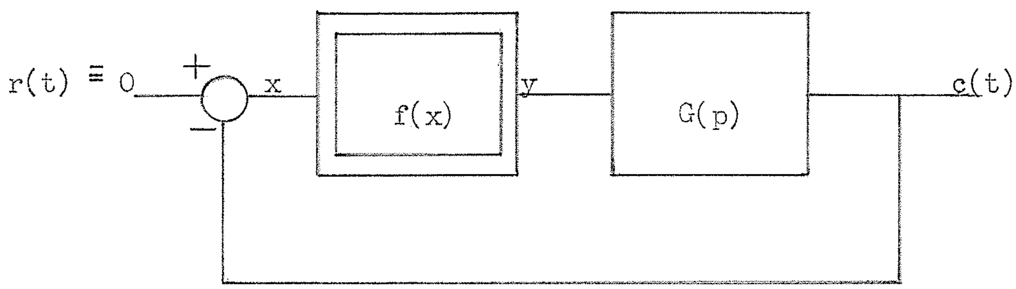


FIGURE 16

CONTROL SYSTEM WITHOUT TIME-LAG

Definition: G(p) is a stable, rational, minimum phase function of order N.

Proof: The system equation in operator form is given by

$$N(p)\left\{f(x)\right\}+D(p)\left\{x\right\} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 121$$

in which N(p), D(p) are polynomials satisfying

$$G(p) = \frac{N(p)}{D(p)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 122.$$

The order of D(p) exceeds that of N(p) and p denotes the operator $\frac{d}{dt}$. The phase equations are given by

$$\left.\begin{aligned} x_1 &= x \\ \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_N &= \frac{d^N x}{dt^N} \end{aligned}\right\} \qquad 123.$$

Let the function, F, be defined by

$$N(p)\left\{f(x)\right\} = F(x, \dot{x}, \ddot{x}, \ldots \frac{d^{N-1}x}{dt^{N-1}}) \qquad 124$$

which becomes

$$N(p)\left\{f(x)\right\} = F(x_1, x_2, \ldots x_N) \qquad 125$$

in terms of the phase variables (Equation (123)). In vector matrix notation, Equation (125) is written as

$$N(p)\left\{f(x)\right\} = F(\underline{x}) \qquad 126$$

in which

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix} \qquad 127.$$

Now D(p) may be expressed as a polynomial with real coefficients, a_o, a_1, a_2 ... a_N, that is by

$$D(p) = \sum_{k=o}^{N-1} a_k p^k + p^N \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ 128$$

in which a_N is taken as unity without loss of generality. By virtue of Equation (126) and Equation (128) the phase equations can now be expressed compactly as

$$\dot{x}_i = x_{i+1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ 129a$$

for i equal to 1, 2, 3, ... N-1 and

$$\dot{x}_N = -(F(\underline{x}) + \sum_{k=o}^{N-1} a_k x_{k+1}) \quad \ldots\ldots\ldots\ldots\ldots\ldots \ 129b$$

for i equal to N.

Suppose $\underline{x}'$ is a (vector) periodic solution to the phase equations, then the replacement of x_i by $x_i' + u_i$, where u_i is the variation in Equation (129a,b) produces

$$\dot{x}_i' + \dot{u}_i = x'_{i+1} + u_{i+1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \ 130a$$

for i equal to 1, 2, 3, ... N-1 and

$$\dot{x}_N' + \dot{u}_N = -\left(F(\underline{x}' + \underline{u}) + \sum_{k=o}^{N-1} a_k (x'_{k+1} + u_{k+1})\right) \ldots\ldots\ldots \ 130b$$

for i equal to N. These reduce to

$$\dot{u}_i = u_{i+1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ 131a$$

for i equal to 1,2,3, ... N-1 and

$$\dot{u}_N = -(\sum_{j=1}^{N} u_j \frac{\partial F}{\partial x_j} (\underline{x}') + \sum_{k=o}^{N-1} a_k u_{k+1}) \quad \ldots\ldots\ldots\ldots \ 131b$$

for i equal to N, when the limit cycle solution is cancelled and the higher order terms in the variations u_1, u_2, u_3, ... u_N are neglected. Equations (131a,b) comprise the (linear) variational equations obtained by a state approach. These phase variational equations are reduced to the single operational variational equation as follows:

From Equation (126), one obtains

$$F(\underline{x}'+\underline{u}) - F(\underline{x}') = N(p)\left\{f(x_1' + u_1)\right\} - N(p)\left\{f(x_1')\right\} \qquad \text{.......... } 132$$

which, from Taylor's Theorem, becomes

$$\sum_{j=1}^{N} u_j \frac{\partial F}{\partial x_j}(\underline{x}') + \text{h.o.t. in } u_1, u_2, \dots u_N = N(p)\left\{u_1 \frac{df}{dx}(x_1') + \text{h.o.t. in } u_1, u_2, \dots u_N\right\} \qquad \text{.... } 133$$

whence

$$\sum_{j=1}^{N} u_j \frac{\partial F}{\partial x_j}(\underline{x}') = N(p)\left\{u_1 \frac{d}{dx} f(x_1')\right\} \qquad \text{.......................... } 134$$

is obtained by equating linear terms in u_1, u_2 ... u_N.

Insertion of the right-hand side of Equation (134) in Equation (131b) produces

$$\dot{u}_N = -(N(p)\left\{u_1 \frac{df}{dx}(x_1')\right\} + \sum_{k=o}^{N-1} a_k u_{k+1}) \qquad \text{..... ....}$$

or

$$\dot{u}_N + \sum_{k=o}^{N-1} a_k u_{k+1} + N(p)\left\{u_1 \frac{df}{dx}(x_1')\right\} = 0 \qquad \text{..................... } 135$$

which, from Equation (128) reduces to

$$D(p)\left\{u_1\right\} + N(p)\left\{u_1 \frac{df}{dx}(x_1')\right\} = 0 \qquad \text{......................... } 136.$$

Since x' is defined as x'_1 and u as u_1, Equation (136) may be recast as

$$N(p)\left\{u\frac{d}{dx}f(x')\right\} + D(p)\left\{u\right\} = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 137$$

which is identical to the variational equation obtained by the operational approach; therefore the theorem is proved.