ADAPTATION AND SATURATION

A DESIGN STUDY

By

PAUL NORMAN HOWELL

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1959

Submitted to the faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
May, 1964

ADAPTATION AND SATURATION

A DESIGN STUDY

Thesis Approved:

Paul A. McCollum
Thesis Adviser

W. A. Blackwell

J. H. Boggs
Dean of the Graduate School

PREFACE

The specific design problem which motivated this study was one encountered by the author while in the employ of one of the leading flight controls manufacturers. Only the symptoms of the problem were known along with most of the system parameters. Specific insight into the probable cause of stability problems came only after the analog simulation mentioned in this thesis.

The presentation of the problem was intended to be as general as possible with the hope that it might be found useful outside the flight controls and inertial navigation field.

The author is indebted to Professor Paul A. McCollum not only for his advice and guidance in the preparation of this thesis but also for the privilege of having been counted among his students in the formative undergraduate years.

A particular expression of thanks is due Mr. Jack M. Walden for his valuable assistance in modifying his root locus digital computer programs to handle the challenging root locus plots associated with the loops in this study. The root locus data presented is the result of the digital computer solutions. Conventional techniques for these plots were impractical for the range of roots encountered.

A word of thanks to my wife Dora is in order for her patience and encouragement throughout my graduate program.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
II.	A BRIEF SURVEY OF ADAPTIVE CONTROL METHODS	6
III.	THE DESIGN PROBLEM	12
	Specifications for a Design	12
	Nonlinear Performance Considerations	15
	Source of Additional Constraints	20
	Applicability of Adaptive Methods	23
IV.	A DUAL-MODE, SELECTIVE FEED-FORWARD LOOP	27
	Complementary Saturation and Dead-Zone	27
	Design of the Auxiliary Path	28
	Performance of Dual-Mode vs. Simple Saturating Loop	35
V.	SUMMARY AND CONCLUSIONS	44
SELECTED BIBLIOGRAPHY		47
APPENDIX A		48
APPENDIX B		53

LIST OF FIGURES

Figure | Page

1. A Simple Control Loop With Describing Function for Saturation Characteristic.................................. 4

2. A Model Reference Passive-Adaptive Control System.......... 7

3. A Simplified Input Signal Adaptive Control System.......... 8

4. A Simplified System Variable Adaptive Control System....... 9

5. A Simplified System Characteristic Adaptive Control System. 10

6. Plant Configuration for the Design Problem................. 13

7. Preliminary Configuration Based on Linear Design Methods... 16

8. Initial Angle Response and Integral of the Squared Error for Linear Loop A.. 16

9. Loop A - Open Loop Gain-Phase Characteristics.............. 18

10. Root Locus Plot for Loop A................................. 19

11a. Initial Angle Response and Integral of Squared Error for Loop A with 30% Saturation................................ 21

11b. Initial Angle Response and Integral of Squared Error for Loop A with 20% Saturation................................ 21

11c. Initial Angle Response and Integral of Squared Error for Loop A with 11% Saturation................................ 22

11d. Initial Angle Response and Integral of Squared Error for Loop A with 10% Saturation................................ 22

12. Simple Model Used for Study of Fundamental Phase Properties.. 28

13. Loop Configuration for the Design of the Auxiliary Channel - Loop B.. 29

14. Loop B - Open Loop Gain-Phase Characteristics.............. 30

15. Root Locus Plot for Loop B................................. 31

Figure | Page

16. Conjectural Root Locus Plot for the Composite Loop......... 32

17. Mechanization of Integral Squared Error Computer........... 33

18. Simulation Diagram for Design Study......................... 34

19a. Comparative Response and Integral of Squared Error for Loop A and Composite Loop for 30% Saturation............. 36

19b. Comparative Response and Integral of Squared Error for Loop A and Composite Loop for 20% Saturation............. 37

19c. Comparative Response and Integral of Squared Error for Loop A and Composite Loop for 15% Saturation............. 38

19d. Comparative Response and Integral of Squared Error for Loop A and Composite Loop for 11% Saturation............. 39

20. Limit Cycles for Two Different Dead-Zones.................. 41

21a. Peak Error Angle vs. Frequency for Loop A for 0.5 sin($2\pi ft$) Input... 43

21b. Peak Error Angle vs. Frequency for Composite Loop for 0.5 sin ($2\pi ft$) Input.. 43

22. Simplified Schematic of a Pendulous Gyro Accelerometer..... 50

23. Block Diagram Showing Motor-Load Transfer Function and Lumped Nonlinearities.................................... 56

CHAPTER I

INTRODUCTION

"By examining a servo system from the error-sensing device to the power stage, an initial acquaintance can be made with non-linearities which, though sometimes overlooked, are always present." (Gille, Pelegrin, Decaulne, 1) Most control system designers are aware of the presence of nonlinearities which are often overlooked and will justly continue to overlook them until the necessity of recognition becomes eminent. Detailed analyses of all the many nonlinear phenomena occurring in typical control systems would doubtless preclude most from the realm of economic feasibility. It may be concluded, therefore, that if a control engineer is to remain competitive and render effective service to those to whom he avails his services, he will incur rigor of analysis only as necessary to effect a satisfactory solution.

It is the intended purpose in this thesis to demonstrate how, with only the basic tools of servo-mechanism theory in addition to a few concepts from the growing area of adaptive control, the control engineer may overcome some formidable problems. Some of these problems involve the meeting of restrictive specifications with express constraints in addition to more subtle constraints imposed by component availability, size, weight, power availability, and even personal foresight, prejudice, or integrity of the designer. How these latter factors mignt enter into the design process will be discussed in Chapter III to a greater extent.

Progress in the control field has been characterized by higher performance requirements, reduced weight, and reduced power requirements. The advent of transistors has led to significant achievements in these areas. The introduction of solid state devices into control circuitry has added at least one perverse factor into the design problem. The transistor, to be specific, operates at voltage levels which are typically much lower than its counterpart in high vacuum thermionic emission devices. The result appears to be that with all the advantages of solid state devices, the control engineer using them must accept a reduced dynamic operating range of voltage when designing controller or compensating networks.

One might propose that the designer merely reduce the noise level by an appropriate order of magnitude and enjoy the same dynamic operating range as with the thermionic devices. Academically, this might at first seem to be a reasonable approach. Express or implied economic constraints, however, often preclude this approach, since the cost of reduction of noise level in electronic devices seems to have an exponential characteristic.

Precise weighting of all the factors or constraints in the design problem is beyond the scope of this thesis. Certain assumptions will be made which, it is hoped, will not significantly alter the conclusions reached.

Subsequent discussion will be based on the premise that a fundamental constraint in the design problem will be the saturation phenomena associated with mechanical stops or limiters, supply voltages, and physical limits on devices such as motor stall torque and peak velocity.

Since motors or some similar transducers often appear in the output stage of a control loop, there are several approaches to the analysis of systems with saturation in the output element. Although the assumption that saturation occurs only in the output element is prevalent in most of the literature, the assumption cannot be made in this study. On the contrary, as will be shown in Chapter III, the significant saturation is far removed from the output element. Clarification should be made, however, that characteristics of the output element significantly contribute to the behavior of the saturating system.

The describing function approach to analysis of lumped nonlinearities is treated by Truxal (2) and others in the literature. Use of the describing function has not been greatly emphasized in the past due probably to lack of rigorous mathematical justification. As with other engineering tools, the rigorous validation has followed later, and now the describing function is receiving renewed attention as an analytical tool. The describing function is essentially a normalized gain schedule of a nonlinear element as a function of input amplitude. Since the describing function may be lumped into a tandem element in the mathematical model of a control system, it becomes apparent that the over-all loop gain is varied according to the value of the describing function. A typical block diagram model of a control loop with a describing function (N) is shown in Figure 1. A brief analysis of a simple saturation characteristic with regard to the definition of transfer gain should be sufficient to justify the idea of an inverse relation between gain and signal level for saturation only. The justification developed by Truxal (2) should be sufficient for the purposes of this study. The root locus method, a powerful analysis and synthesis tool introduced by Evans (3), provides a method to determine the closed loop characteristic roots (poles and zeros) from the open loop characteristic roots as a function of loop gain. No qualification pertaining to the method of variation of the gain is given with the justification of the root locus method. It shall be assumed, therefore, that gain variations resulting from saturation shall be fully as valid as variations derived in any other manner.

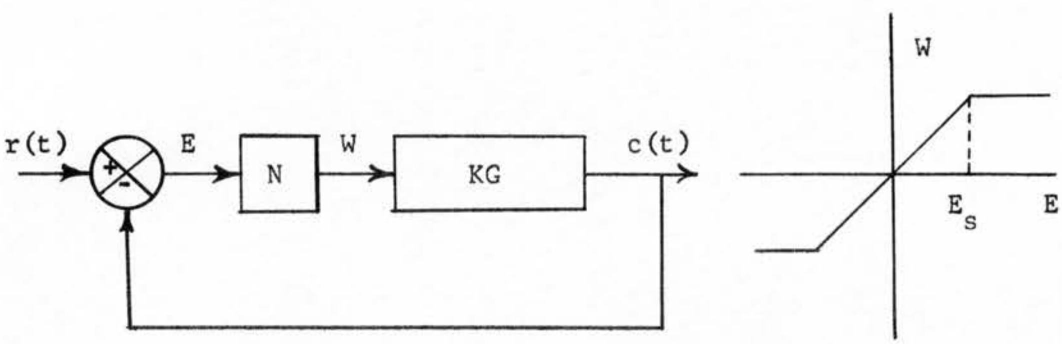

Figure 1. A Simple Control Loop With Describing Function for Saturation Characteristic

Since variations of closed loop poles of a control system may be effected by signal level in the loop indirectly through the saturation phenomena, extreme variations of the roots of a "conditionally stable" system might result in a rendering of the system into an unstable and divergent state. Motivated by the knowledge that destructive instability can and has occurred as a result of the saturation characteristics or more generally to the nonlinear characteristics of the type to be described in Chapter III, a survey of current design philosophies was made in search for a method to overcome such tendencies.

By reasoning that saturation actually effects pole shifts, one particular design philosophy immediately appears to be applicable to this type problem. Adaptive control, through its assumption of variational or inexactly known plant parameters, appears to be a likely area in which to seek a solution to the saturation dilemma. In the next chapter, several basic adaptive control methods will be described and discussed briefly. With this background, a more perceptive view of the design problem may be possible.

CHAPTER II

A BRIEF SURVEY OF ADAPTIVE CONTROL METHODS

In recent years a new area of control engineering has emerged bearing the name adaptive control. Much controversy exists at this writing as to the exclusiveness of the set of problems (or their respective solutions) to which this label should be applied. The term "adaptation" according to Harris (4), was borrowed from the biological sciences where it is defined as a "modification of an animal or plant (or of its parts or organs) fitting it more perfectly for existence under the conditions of its environment."[1] Thus, if we extend this definition into the control field, we mignt assume that an "adaptive control system" somehow effected a change in its parameters or performance in response to its environment. We might also extend the meaning of environment, in the case of control systems, to include all ambient conditions in addition to input signals, output loading, intermediate injection of noise or other corrupting signals as well as variations in performance criteria.

Definitions of "adaptive control" are almost as numerous as are authors in the control field. One in particular seems to fit the study presented. "An adaptive system is any physical system designed with an adaptive point of view." (Truxal,5) In light of the following defini-

[1] "Adaptation," Webster's New Collegiate Dictionary, G. & C. Merriam Co., Publishers, (Springfield, Massachusetts, 1956), p.10.

tions and discussion, the design presented in Chapter IV may have no other basis for such a classification.

<u>Passive Adaptation</u>

The passive-adaptive system is one in which the desired performance or approach to the desired performance is effected without active adjustment of system parameters. Systems such as "ordinary feedback" systems as described by Horowitz (6) come into this category. In addition to all the classical feedback or compensating techniques, the idea of the model reference technique in which no active parameter adjustment is made, can also be classified as passive-adaptive. Lang and Ham (7) show that the linear passive-adaptive system can be represented by a linear prefilter at the control system input. Horowitz (6) further shows that so long as the system designer retains two degrees of freedom in his design, any arbitrary insensitivity of a particular parameter may be effected.

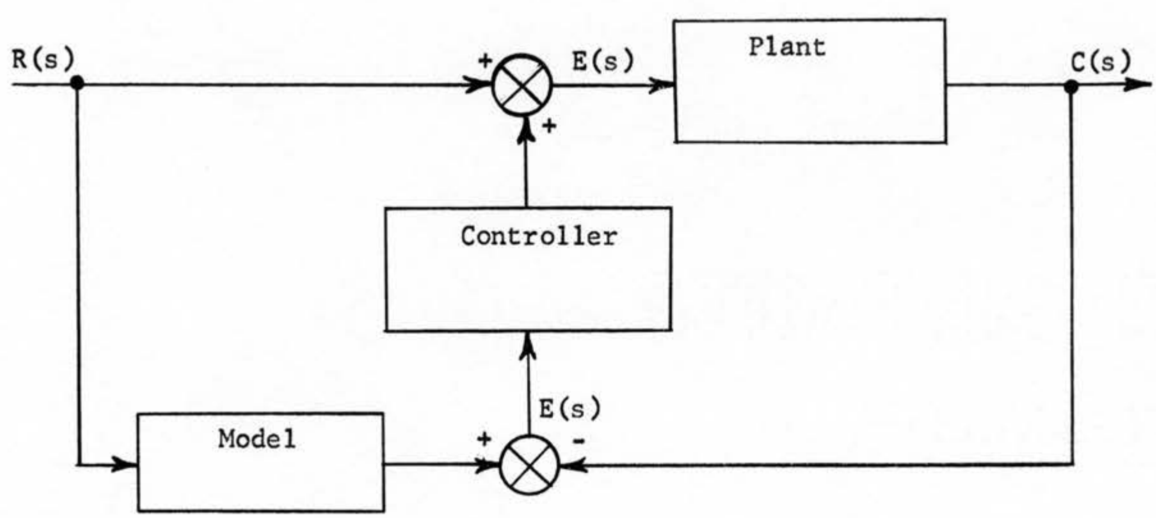

Figure 2. A Model Reference Passive-Adaptive Control System

Input Signal Adaptation

Those systems which adjust their parameters according to the nature of the input signal may be classified as input signal adaptive. They are comprised essentially of some identification component which derives information from the input signal, which in turn is used to actuate a controller which changes the plant parameters. This type system evidently assumes a priori knowledge of the nature of possible input signals and a program for desired performance exists for each type. This type system, however, does not seem to consider the possible variations in plant parameters unless they are associated uniquely with an input situation.

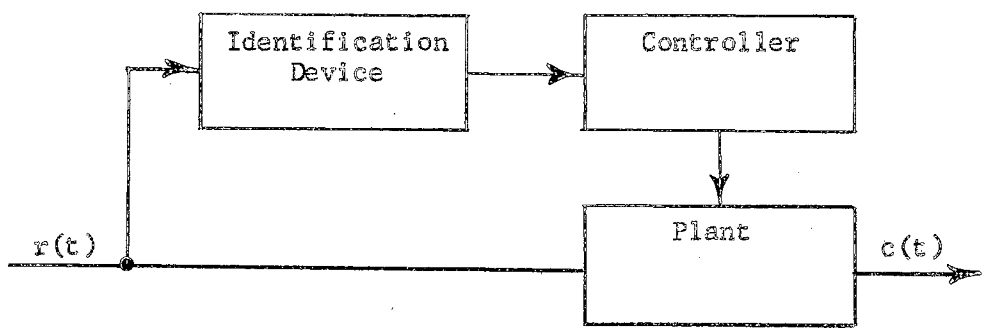

Figure 3. A Simplified Input Signal Adaptive Control System

Extremum Adaptation

The control system which consistently seeks the extremum for a predetermined criterion is of considerable interest to many control engineers. This type system may in itself be responsible for bringing practitioners of various disciplines such as economics, mathematics, and engineering, together in solving or more primarily defining the various optimization problems amenable to techniques such as dynamic programming. Bellman and Dreyfus (8) describe several basic problems in this area.

System Variable Adaptation

Systems of this type might perform some analysis of its output, input, error, or any combination thereof. The analysis possibly determines magnitudes of variables or derivatives of the variables, and adjusts plant or feedback parameters. Randomly programmed and logical determination of variables to be analyzed seem to be equally permissable. A simplified model of a system variable adaptation scheme is shown in Figure 4. Several variations on this method are documented by Aseltine, Mancini, and Sarture (9) but the example illustrated should be sufficient for purposes of this study.

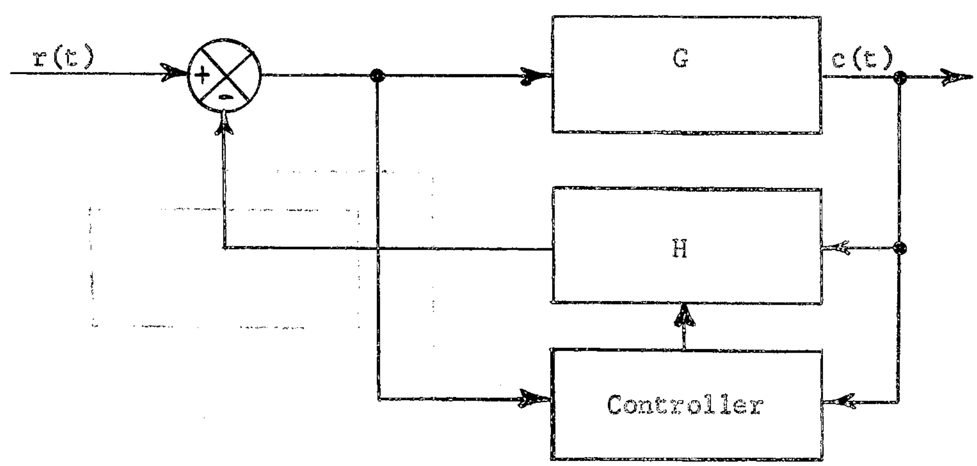

Figure 4. A Simplified System Variable Adaptive Control System

System Characteristic Adaptation

Probably the most discussed class of adaptive control systems are those which effect some parameter adjustment in accordance with or in response to some system or plant identification process. Some of these systems effect plant identification by periodically taking the control system "off line," applying a signal to the input and performing some type of convolution process with the output to definitely identify the plant function. Once identified, a comparison may be made with the desired plant function and a controller actuated to effect changes to reduce discrepencies between desired and actual plant function. Other systems, including at least one currently marketed autopilot system, performs the system identification through the use of a "dither" signal of low amplitude applied to some point in the system. This dither signal is of such amplitude and frequency so as to achieve the desired identification without adversely affecting the performance characteristics of the system. This approach overcomes the problem of "off line" time which in many cases may be absolutely prohibited.

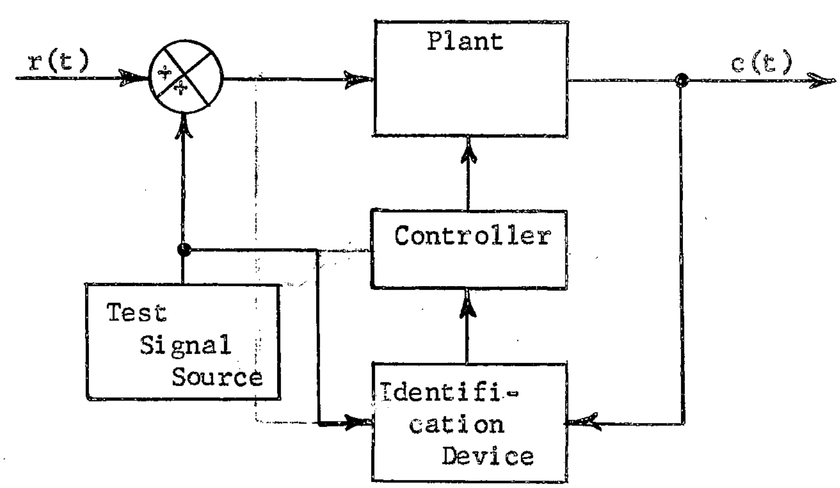

Figure 5. A Simplified System Characteristic Adaptive Control System

Variations on this method of adaptive control are too numerous to mention here, but the basic concept of identification, and active adjustment of parameters are the two prime factors pertaining to this type system to which we will refer in the later chapters.

Adaptation by Learning

Systems involving memory and learning capabilities are not classified separately by Aseltine, Mancini, and Sarture (9) but a separate category seems appropriate. Gibson (10) and others conjecture on the unprogrammed, arbitrary memory and decision network in which paths or bonds are retained or destroyed according to positive or negative reactions to the originally arbitrary decisions. Here again exists a crossing of disciplinary boundaries into areas previously encountered only by psychologists and sociologists. Mowrer (11) describes the positive stimulus-response (S-R) bond and the erasure of such bonds. He further introduces the idea of conflict (or negative bond) and fear motivation which, so far, has not been treated to any significant extent in control literature. Further elaboration on the learning systems will be omitted due to lack of development in this area. These systems are of considerable academic interest, however, and probably merit much of the effort expended upon their study.

It should appear obvious upon encountering the design problem detailed in the next chapter, that some of the adaptive methods above could not receive serious consideration. Their inclusion in this discussion however, seems desirable for background.

CHAPTER III

THE DESIGN PROBLEM

The design problem presented in this chapter might be considered typical to the designer of inertial instrumentation. The primary plant, in this case, might consist of a precision gyroscope whose scale factor stability, reliability, size, weight, and drift characteristics have contributed to its specification as a primary component in a control system. From the control system designer's point of view, the primary plant is absolutely fixed and alternate primary plants are precluded.

Specifications for a Design

For a given "off the shelf" plant such as shown in Figure 6, design appropriate amplifiers, demodulators, compensation networks, and controller (or driver) satisfying the following requirements and constraints:

1. System band-width less than 30 cps.
2. Threshold requirements require 16,150 volts/radian minimum gain (steady state) through the designed portion of the loop.
3. System inputs up to 0.5 rad/sec equivalent $\leq$ 1 cps.
4. Inputs of reduced amplitude above 1 cps.
5. No steady state error can be tolerated (output will be integrated).
6. Accuracy of system must remain dependent upon the primary plant.
7. Dynamic error angle (primary plant output) must remain less than

its open loop maximum to avoid destruction of the primary plant.

8. Transistor circuitry only (weight and power requirements).

9. Primary plant gain factor given K_1 = 7 rad/rad.

10. Primary plant characteristic time given t_1 = 0.0067 sec.

11. Secondary plant gain factor given K_2 = 8.85 rad/sec/volt.

12. Secondary plant characteristic time given by (1.0 + .025 R_o) sec (where R_o is the output resistance of the controller or driver element).

13. One integration occurs in the plant, either in the primary or secondary, depending on the nature of the output assumed.

14. Maximum output of controller or driver shall be $\pm$20 volts d.c.

The above specifications, of course, do not include all the considerations involved in a design of this type, but the essential limitations are included.

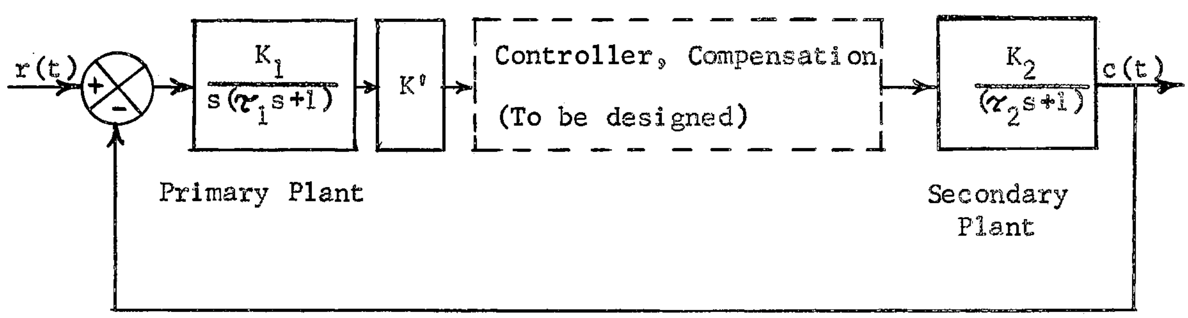

Figure 6. Plant Configuration for the Design Problem

The secondary plant (a platform and motor in the case of a typical inertial instrument) is subject to a finite number of choices if, as is often the case, the secondary plant is a purchased part. For this discussion, the secondary plant shall be considered fixed also as the specifications imply. The nature of the transfer function for the primary plant output error angle is fixed for the primary plant as an alternating current signal voltage. The magnitude of the transfer function is variable within certain limits and adjustment may be made by the designer by varying the magnitude of the reference voltage to the pickoff. The primary consideration in determination of the pickoff reference level is usually reduction of quadrature and harmonics and is, therefore, fixed for a given signal amplifier. The remaining factor in the plant, adjustable by the designer, is the time constant of the secondary plant (or motor). How the driving source affects the time constant of a d.c. motor is detailed in Appendix B.

With the output of the primary plant being an a.c. signal and the input requirements of the secondary plant being direct current, the distribution of the gain with respect to the required phase-sensitive demodulator becomes significant. At low signal levels phase sensitive demodulators suffer from unbalance due to diode characteristic non-uniformity, diode temperature characteristics and aging as well as harmonics in the reference and signal voltages. Practical limits on quadrature voltage and harnomics depend primarily on the design of the primary plant, and secondarily on the designer specified minimization procedures that become part of the assembly process. It goes almost without saying that harmonic content of the signal is directly dependent on the harmonic content of the reference voltage which must also be specified by the designer.

At this point the designer might resort to the use of a Nichols chart or similar design procedure to arrive at the open loop characteristics typified by the loop shown in Figure 7. The step response or loop closure from an initial angle along with a plot of the integral of the squared error is shown in Figure 8. This and subsequent responses were obtained from the analog computer simulation described later in this chapter.

The response of this linear loop is certainly satisfactory by most standards and seems to meet the specifications in every detail.

Nonlinear Performance Considerations

As discussed previously in relation to describing functions, both dead-zone and saturation can affect the gain sufficiently enougth to cause instability in a conditionally stable loop of the type in Figure 7. Subsequent reference to this loop will be made as Loop A or the loop designed by linear considerations, since it represents the linear or small signal solution to the original design problem. An open loop gain-phase plot for Loop A is shown in Figure 9 while the root locus plot for the same loop is presented in Figure 10. For the purposes of this study, an initial assumption will be made that gain variations are caused only by saturation. Later, the design result will be evaluated considering dead-zone also. Saturation levels as a result may be exaggerated.

It has been noted that typically the amplification stage prior to the demodulator is designed so that the gain after the demodulator is insufficient to saturate the output member by unbalance voltages occurring at the demodulator output. The gain after the demodulator, it should be remembered, typically contains the attenuation of the d.c. compensation

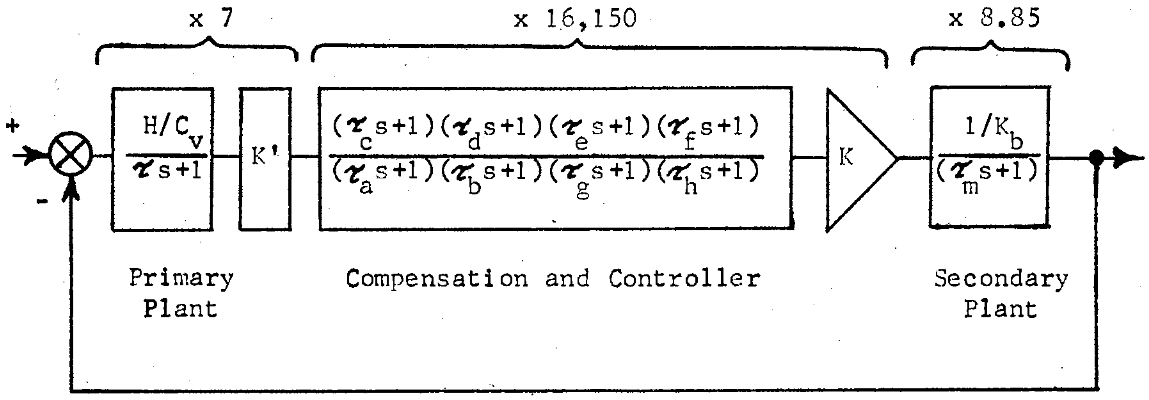

Figure 7. Preliminary Configuration Based on Linear Design Methods

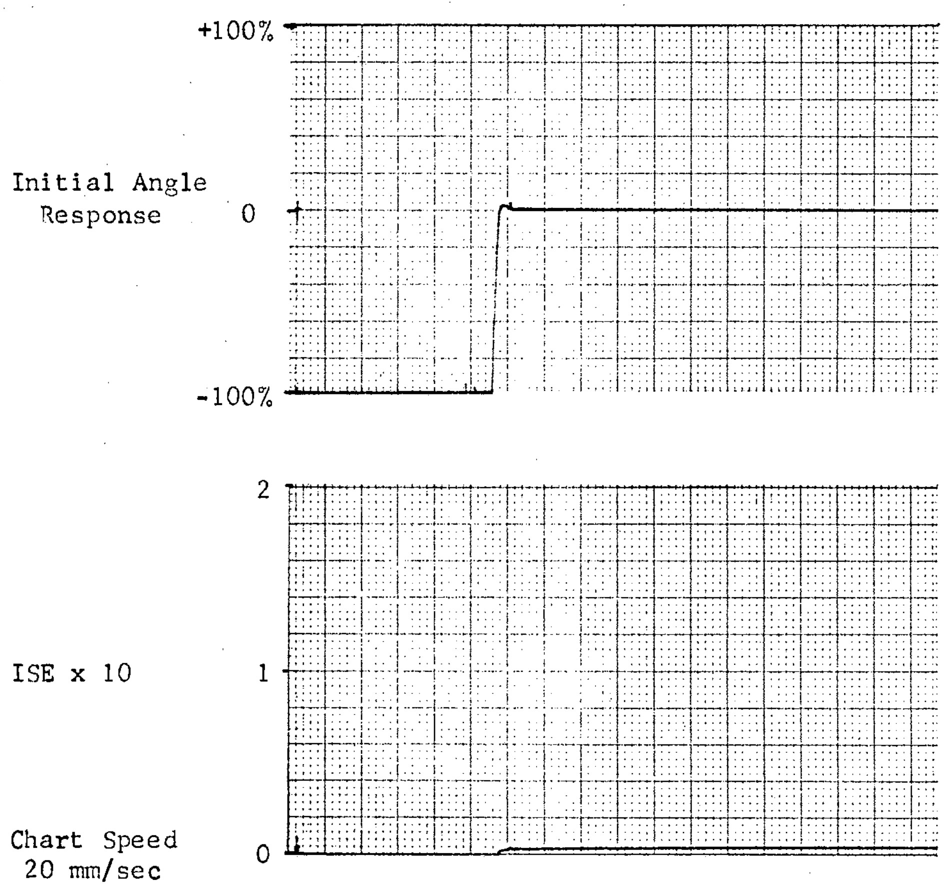

Figure 8. Initial Angle Response and Integral of the Squared Error (ISE) for Linear Loop A

stages.

Although open loop considerations would indicate that saturation would occur in the output member, actually the saturation associated with typical inputs beyond a given magnitude occurs not at the output element, but at the first amplification stage of the typical "designed" portion of the loop. One might, at first, expect the primary saturation tendency to appear at or near the output element of the loop. Further investigation verified by analog computer simulation shows that the closed loop primary saturation did, in fact, occur at the first amplifier of the typical "designed" portion simulated. One might attribute this phenomena to the formidable attenuation characteristic of the compensation network at normal system input frequencies.

Figure 9 shows the gain and phase characteristics for the loop designed by linear considerations. A band of gain values is also indicated which shall be assumed to represent the gain levels for the degrees of saturation discussed. For typical operating conditions with the loop originally designed, no particular adverse characteristics would be expected.

As saturation is introduced into the first controller stage, the critical frequency of the loop is lowered, until, for a saturation level corresponding to about ten per cent or minus twenty decibels, the phase margin becomes zero. If other nonlinearities were considered, the critical frequency would be higher. Figure 11 shows the step (initial angle) response for several saturation values of Loop A. As may be seen, for simple saturation of about ten per cent, the loop becomes divergent and exceeds the boundary specified for destruction.

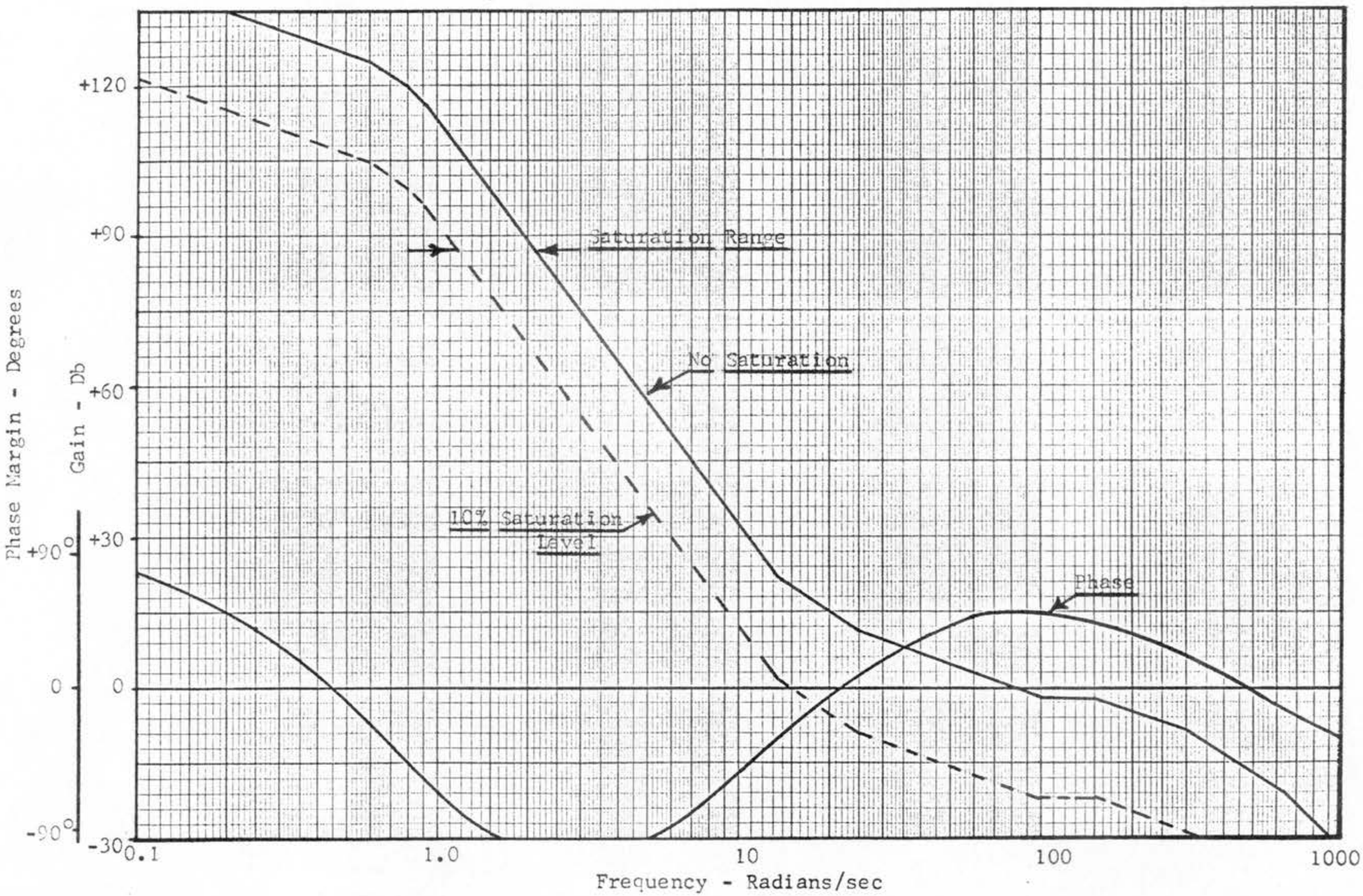

Figure 9. Loop A - Open Loop Gain-Phase Characteristics

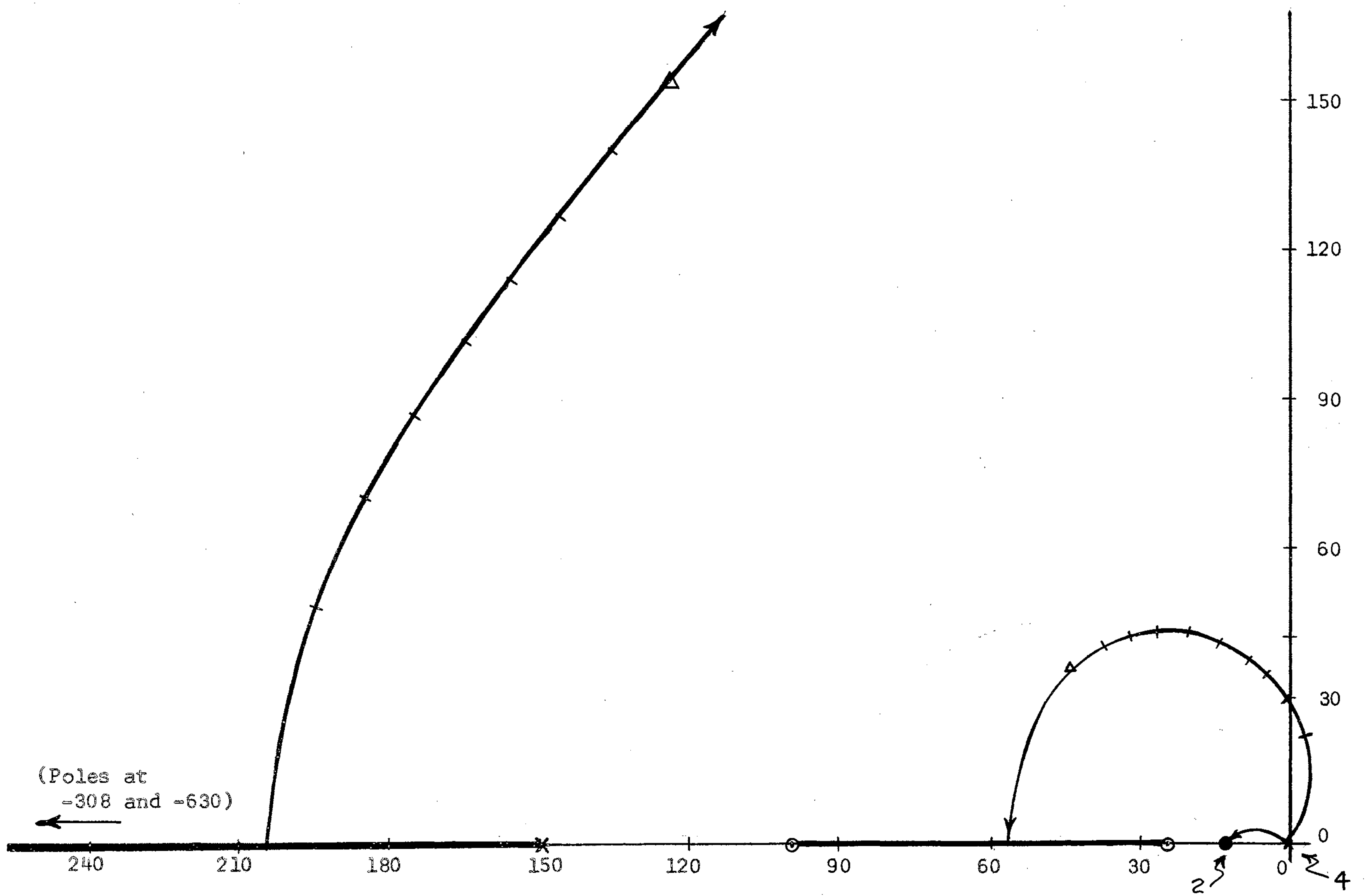

Figure 10. Root Locus Plot for Loop A

Source of Additional Constraints

It seems that even though expressed specifications might not rule out such behavior for system inputs beyond the expected level that the subtle constraint possibly would now be considered.

The design has met the express performance requirements and further has met them within the design constraints. However, with full knowledge that a control system would operate satisfactorily at rated input levels but would become absolutely and destructively unstable at input levels only slightly more than rated levels, the responsible control system designer would probably consider his solution to the design problem inadequate in light of his own conscience.

Presently, most control systems of the type discussed in this study are used in unmanned and as yet non-operational military weapon systems. In the case of system input exceeding that specified by contract, a total failure of a subsystem such as the one with which this study is concerned, would result in the usual marginal weapons test identified with the spectacular fiery explosion at the hand of a range safety officer in charge of a weapon systems launch area. In such a hypothetical case as just mentioned, probably only a few million dollars of defense budget were involved. It seems relatively easy to charge off such amounts to experience. Such is not the case with manned vehicles, however, which tend to include the less accurate but more reliable navigation and control subsystems.

Inertial navigation systems for both manned spacecraft such as those for lunar excursions and commercial aircraft are presently being designed. Inertial navigation facilities for commercial airlines are sought to overcome the present problems of limited visibility and the associated

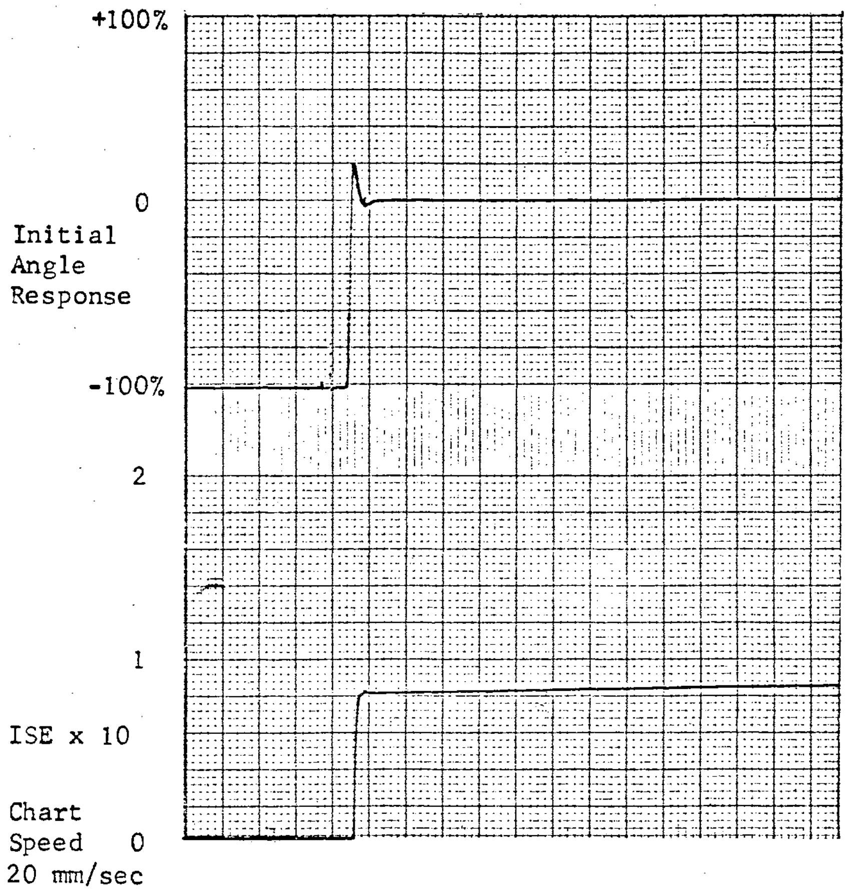

Figure 11a. Initial Angle Response and Integral of Squared Error (ISE) for Loop A With 30% Saturation

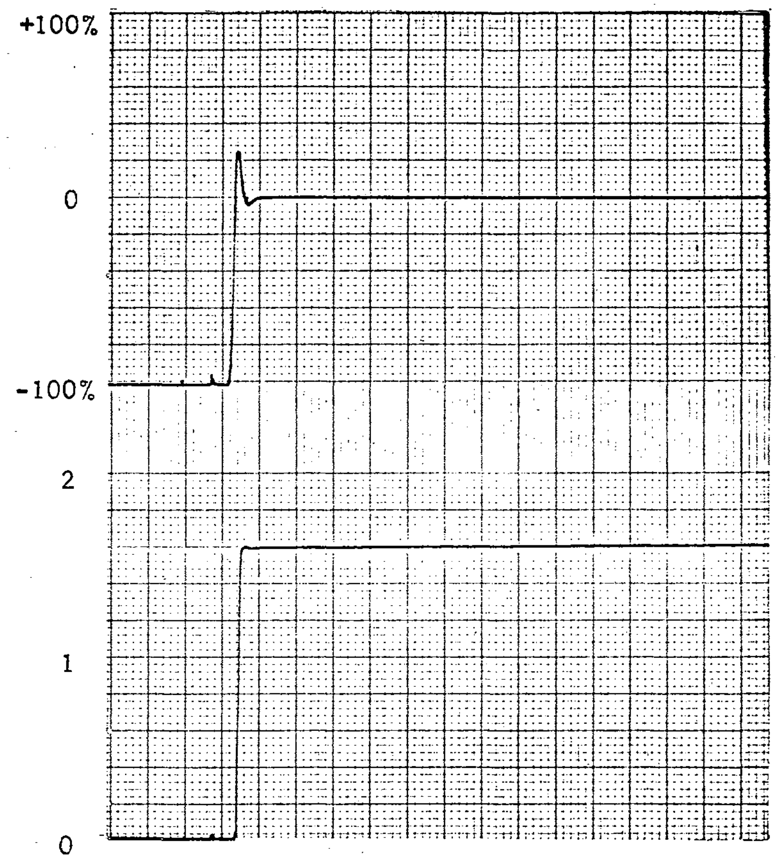

Figure 11b. Initial Angle Response and Integral of Squared Error (ISE) for Loop A With 20% Saturation

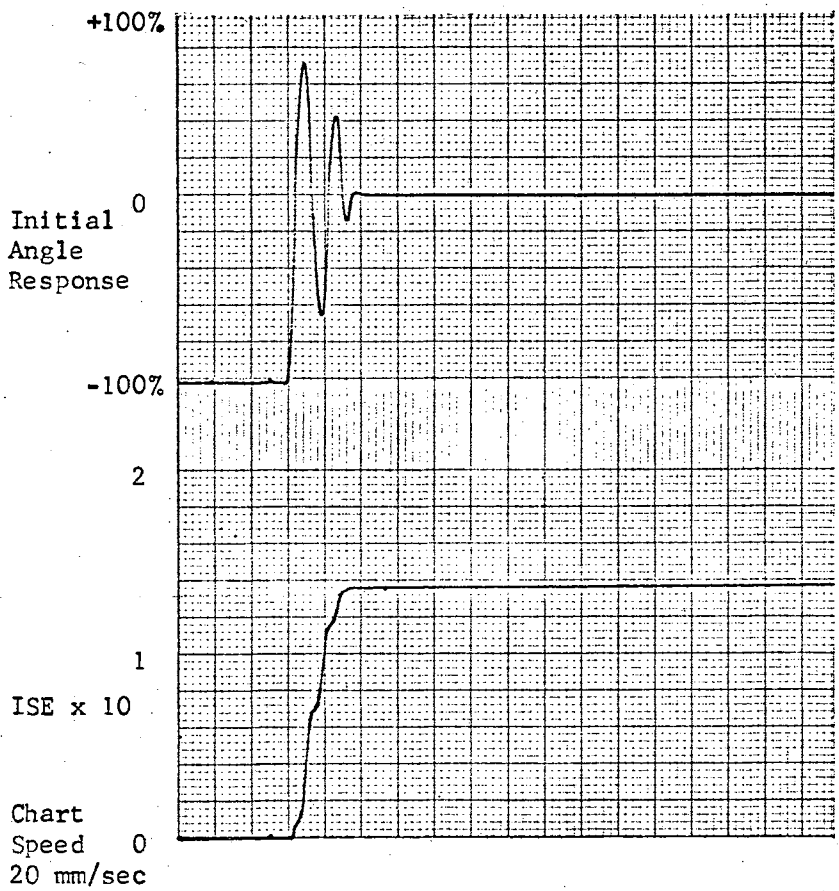

Figure 11c. Initial Angle Response and Integral of Squared Error (ISE) for Loop A With 11% Saturation

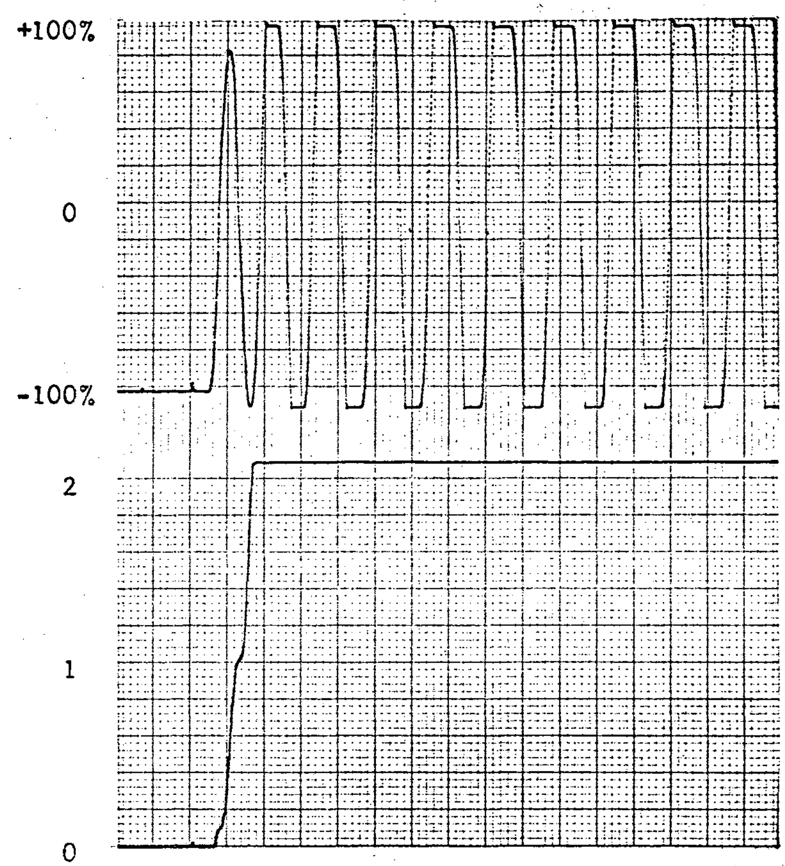

Figure 11d. Initial Angle Response and Integral of Squared Error (ISE) for Loop A With 10% Saturation

shutdown of major air traffic terminals. The improved performance accuracy for lunar excursions have required consideration of control loops such as the one considered in this study.

If a control system designer were faced with the same dilemma as before, he most surely would classify his product as inadequate where the safety of human life was in question. If in such a case, the requirement were not already expressed, he would probably add the following performance requirement: for system inputs beyond the normal levels, the system will perform at reduced accuracy until the excessive level is removed, and at that time will resume operation at the required performance level. It is addition of constraints such as these that might be referred to as imposed by personal bias or integrity. For such an extreme case, the role of the subtle constraint appears almost oversimplified.

Applicability of Adaptive Methods

Since the adaptive control field seems to have sprung up on the idea of ill-defined or variational plant parameters, the methods presented in Chapter II will be reviewed for applicability to this problem.

Input Signal Adaptation

This sounds at first to be a likely approach to the problem at hand since the saturation problem is a result of the input signal exceeding a particular limiting value. If the input signal adaptation scheme is examined more closely, however, the implication appears that some parameter is available for adjustment after the input signal has been examined and found to be unsatisfactory for the loop in its present status. At least two implications are counter to the needs of the saturating system. First, a priori knowledge of the nature of possible input signals seems to rule out the possibility of unpredictable inputs. Second, although not expressed, there seems to be a finite time implied for the process of signal identification and parameter adjustment. Further, there was not available from the outset of the problem, any external method to adjust plant parameters. It shall be assumed that input signal adaptation is more applicable in the case of slowly varying parameters and is of little use in rapid saturation phenomena.

Extremum Adaptation

This method appears further from applicability than the last approach. The original system probably gives maximum performance under normal operating conditions, and for the short times when normal operation is exceeded, it is not the purpose of the system to meet any particular extremum other than to remain stable. Again this method apparently assumes slowly varying parameters from the outset and consequently is inapplicable to the suddenly saturating system.

System Variable Adaptation

This method, currently popular in the literature, fails in several ways to meet the requirements of the saturating system under study. First, it would be absolutely out of the question to assume that any benefit could be derived by taking the system "off line" for a period of time to adjust its parameters, even if they were adjustable. The same limiting considerations apply as with the previous case. Test signals or "dither" applied during on line operation not only would not solve the problem of sudden unpredictable input but also would likely contribute to errors due to "anisoelasticity," or a sort of mechanical rectification. Such considerations are beyond the scope of this study and will not be pursued further. It should suffice to say that, as before, certain identification or correlation time is inherent in this method and thereby becomes inapplicable to the problem at hand.

<u>Adaptation by Learning</u>

An overwhelming shortcoming of such a system as this is its absence from the realm of current synthesis techniques. It could hardly be considered useful otherwise since saturation (hopefully) is highly unusual and as in the case of input signal adaptation, a consideration of the saturating signal from a normal performance point of view would detract from the performance of normal signals. Thus the learning system, as each system considered, fails to offer a satisfactory approach to sudden or stochastic saturation tendencies.

<u>Passive Adaptation</u>

This adaptive method including "classical feedback" and its variations, possesses the unique characteristic of performing at the same rate (or concurrently) with the original systems. This may be verified fairly simply by a simple block diagram manipulation of the passive-adaptive system using a model reference. Lang and Ham (7) show convincingly that the passive-adaptive approach to system design is essentially equivalent to the addition of a prefilter to the original loop. Horowitz (6) has shown further that, given two degrees of freedom, the passive-adaptive approach to system design is probably fully as effective as any of the other adaptive approaches.

One of the first papers to promote the design philosphy that later became known as adaptive control was that by Lang and Ham (7) mentioned previously. It was through this article and the knowledge of the existing problem, that the non-switching, dual-mode solution described in

Chapter IV was motivated. Although the method presented does not appear in the literature, it was classified as passive-adaptive more from the design approach than from other considerations.

CHAPTER IV

THE DUAL-MODE OR SELECTIVE FEED-FORWARD LOOP

Since, as Lang and Ham (7) show, the model reference passive-adaptive control can be represented by an equivalent prefilter, it seemed that an equivalent nonlinear prefilter might be effected also. The basic assumption of necessity being that the error, not the input, was available to use for compensation. Obviously a straightforward equivalence to some compensator in the loop for a nonlinear prefilter is not easily obtained. It is at this point that a search for exact equivalence is abandoned in favor of a more intuitive approach.

Complementary Saturation and Dead-Zone

Studies performed on the analog computer have indicated that for a complementary saturation and dead-zone as in Figure 12, several interesting observations can be made. If the output of the two nonlinearities of Figure 12 are fed into simple lead and lag networks respectively, the linear summation of the network outputs has a peculiar characteristic. If the respective magnitudes of the fundamental components are equal, the linear summation will have zero net phase shift. This may be verified by sketching a simple phasor diagram of two equally leading and lagging voltages and adding them vectorially. Further, as may be verified by the same method, when the voltages are not the same magnitude but remain equally leading and lagging in phase, the output of the summing point will have a phase characteristic tending toward that of the larger signal. Thus the phase of the output is varied both by the phase and the amplitude of the inputs.

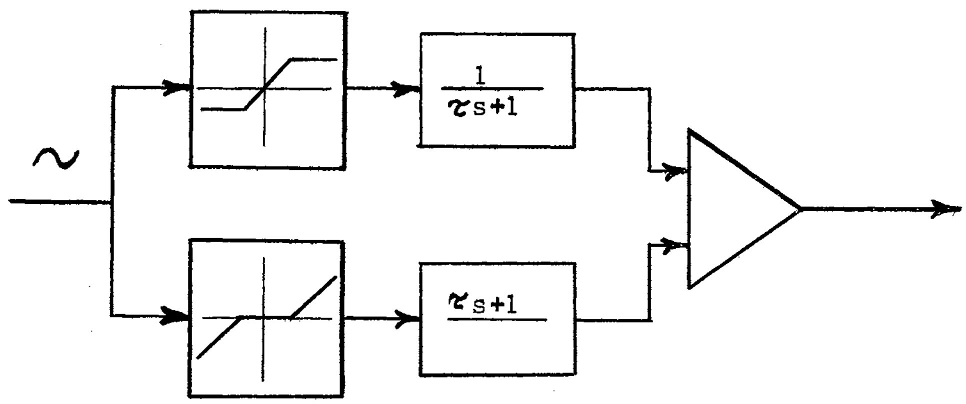

Figure 12. Simple Model Used for Study of Fundamental Phase Properties

Design of the Auxiliary Path

As has been shown previously, the critical saturation level in Loop A occurs when the input level is approximately ten times the saturation level. This critical saturation and its associated gain correspond to the neighborhood of a net 180° phase shift on a Bode gain-phase plot. (See Figure 9.)

With this idea in mind, Loop B was designed to have an equivalent gain of one tenth that of Loop A. Futher, Loop B was designed to have a maximum phase margin at precisely the frequency at which the phase margin of Loop A was zero. In this way, using the complementary saturation and dead-zone discussed previously, a linear combination of the signals of Loop A and Loop B just prior to the secondary plant function yields an equivalent phase margin of one half that of Loop B. At saturation levels below the maximum saturation level of Loop A, the phase margin of Loop A increases while the phase margin of Loop B diminishes. Thus, when the signal level barely exceeds the saturation level of Loop A, Loop B represents an unstable contribution to the total loop performance but of insignificant magnitude.

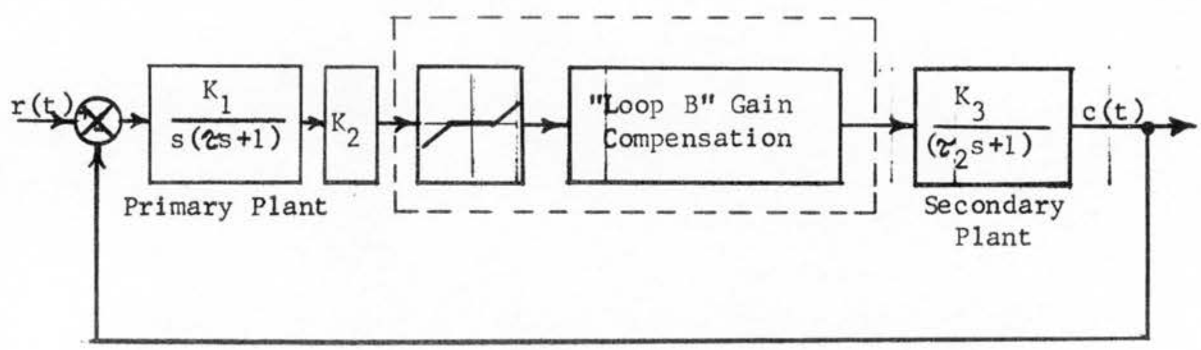

Figure 13. Loop Configuration for the Design of the Auxiliary Channel (Loop B)

Figure 9 shows open loop gain and phase for the preliminary design (Loop A). Figure 14 shows the corresponding characteristics of the auxiliary loop (Loop B). It should be noted that the auxiliary loop in this case has precisely the same pole and zero configuration and differs only in the magnitude of the poles and zeros and of the gain.

Since the poles and zeros of the two loops are ordered similarly one might conjecture that some set of paths could be defined between the root positions on each of the root locus plots such that for any saturation level, the instantaneous location of closed loop poles and

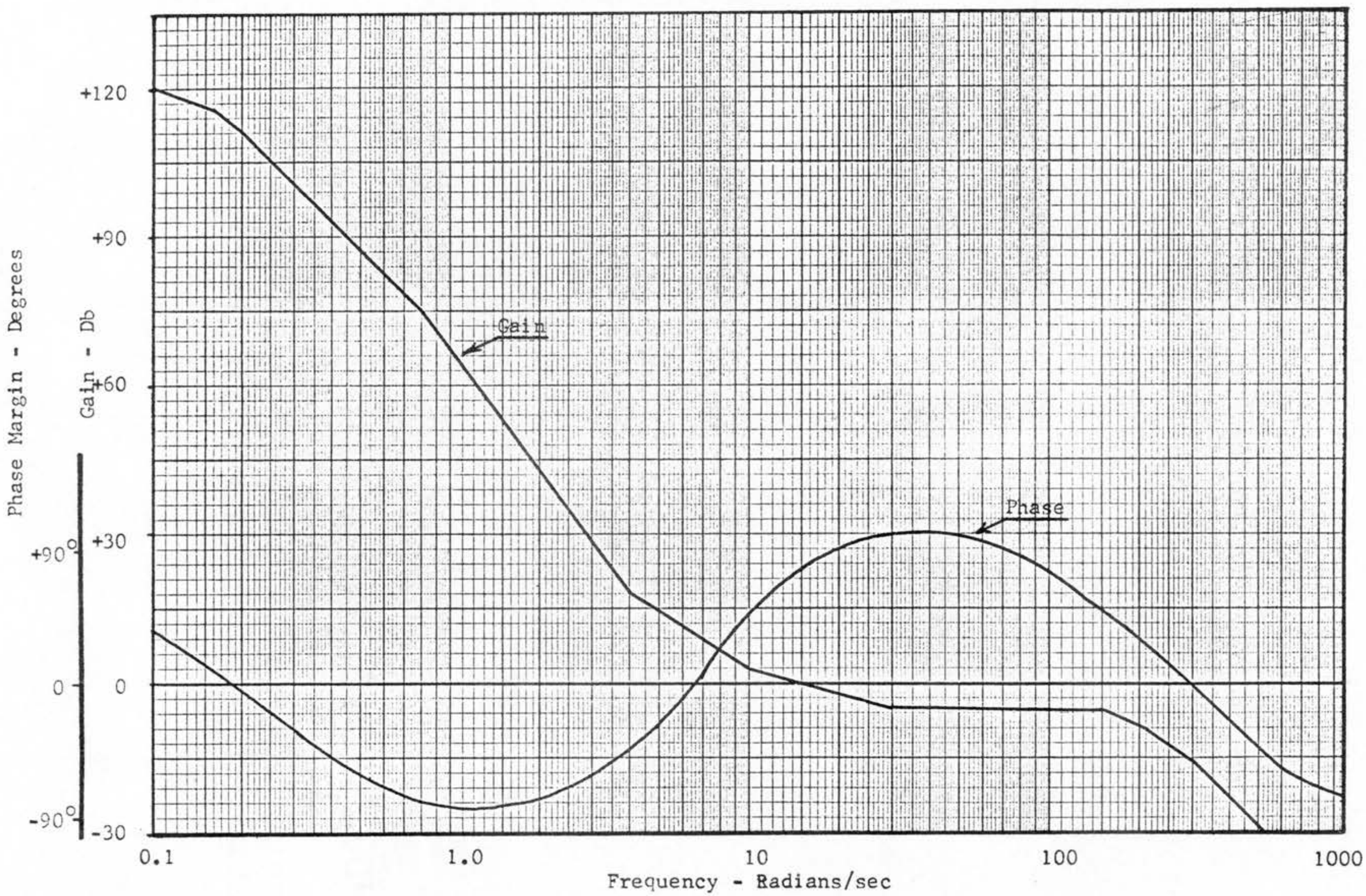

Figure 14. Loop B - Open Loop Gain-Phase Characteristics

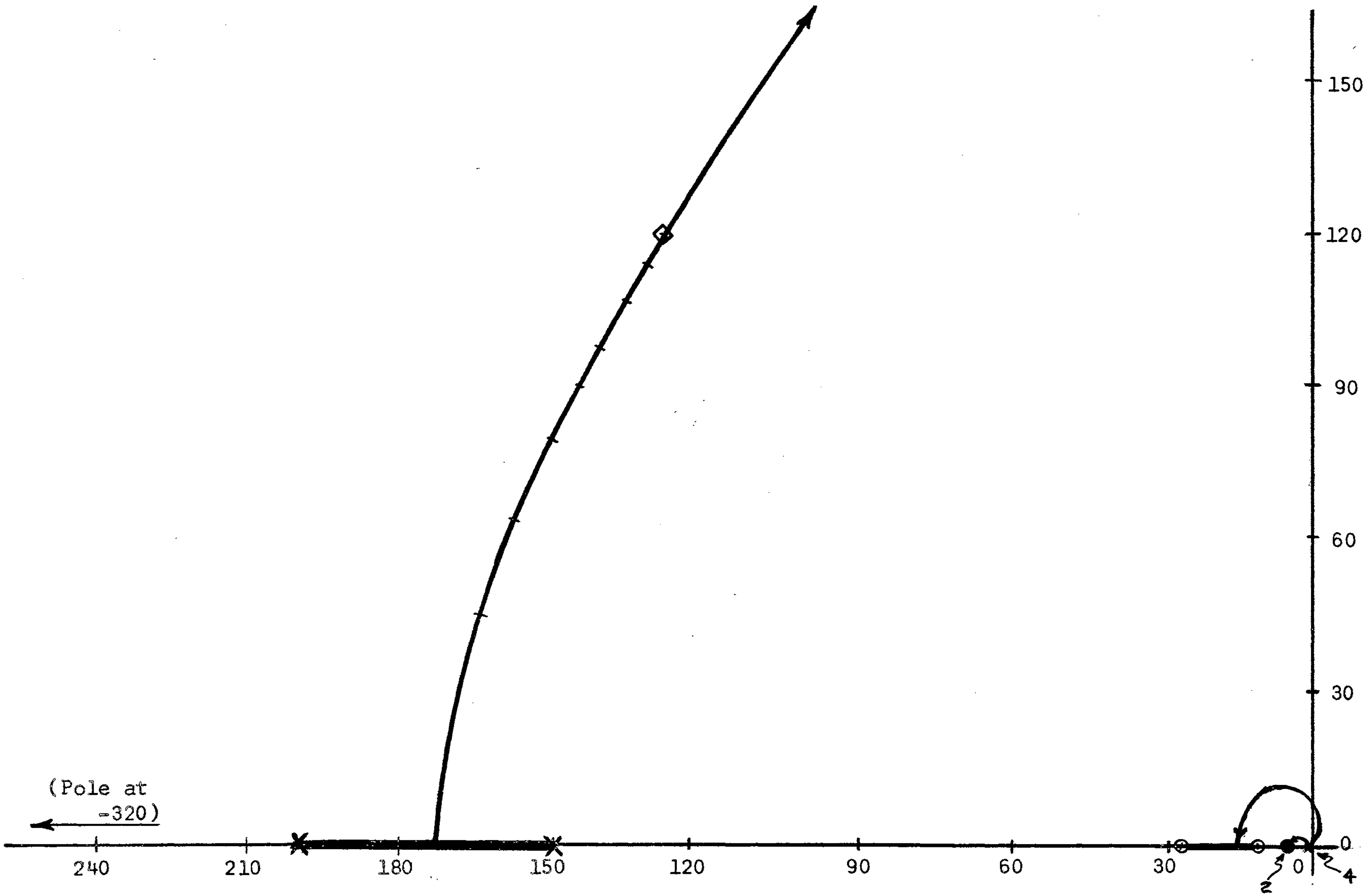

Figure 15. Root Locus Plot for Loop B

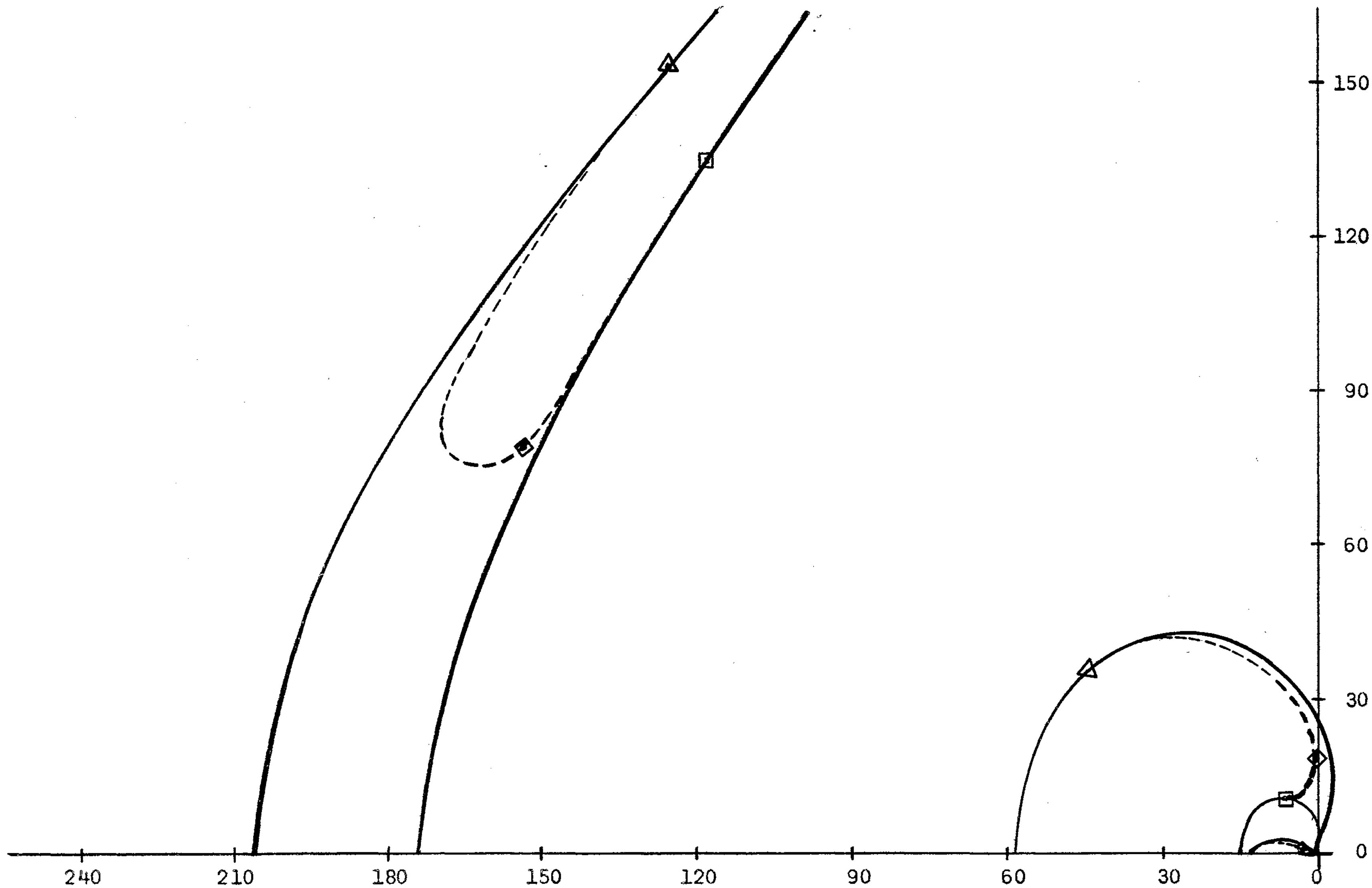

Figure 16. Conjectural Root Locus Plot for the Composite Loop

zeros could be determined for the composite (dual-mode) system. The upper left quadrant of the root loci for Loop A is shown in Figure 10, and for Loop B in Figure 15. In Figure 16, both loci are plotted and in addition, the conjectural paths that might represent the combined system root loci. Further argument for sustantiating this idea will not be presented in this discussion, but since the idea entered into the motivation for the dual-mode loop, it was included for completeness.

An analog simulation of the entire loop was mechanized on the analog computer. A diagram of the simulation is shown in Figure 18.

As may be seen from the analog computer mechanization, the dead-zone associated with Loop B was the complement of the saturation of Loop A. This method reduces the possibility of variations in response data due to inexact duplication of dead-zone settings for given saturation levels. The data gathered in this way and with mechanization mentioned was very consistent from test to test.

In addition to the mechanization of the composite loop shown in Figure 18, an additional integral of the squared error circuit was provided. Figure 17 shows in block form the arrangement for obtaining the integral squared error (ISE). No particular effort was given to dimensioning the ISE since its use was merely to compare the various loop combinations and saturation levels. Particular attention should be given however to the magnitudes on the graphs in addition to the multiplier to be used with the graph.

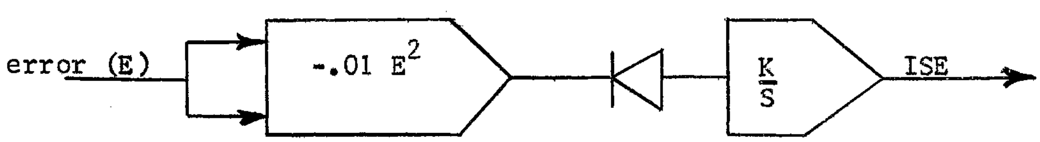

Figure 17. Mechanization of Integral Squared Error Computer

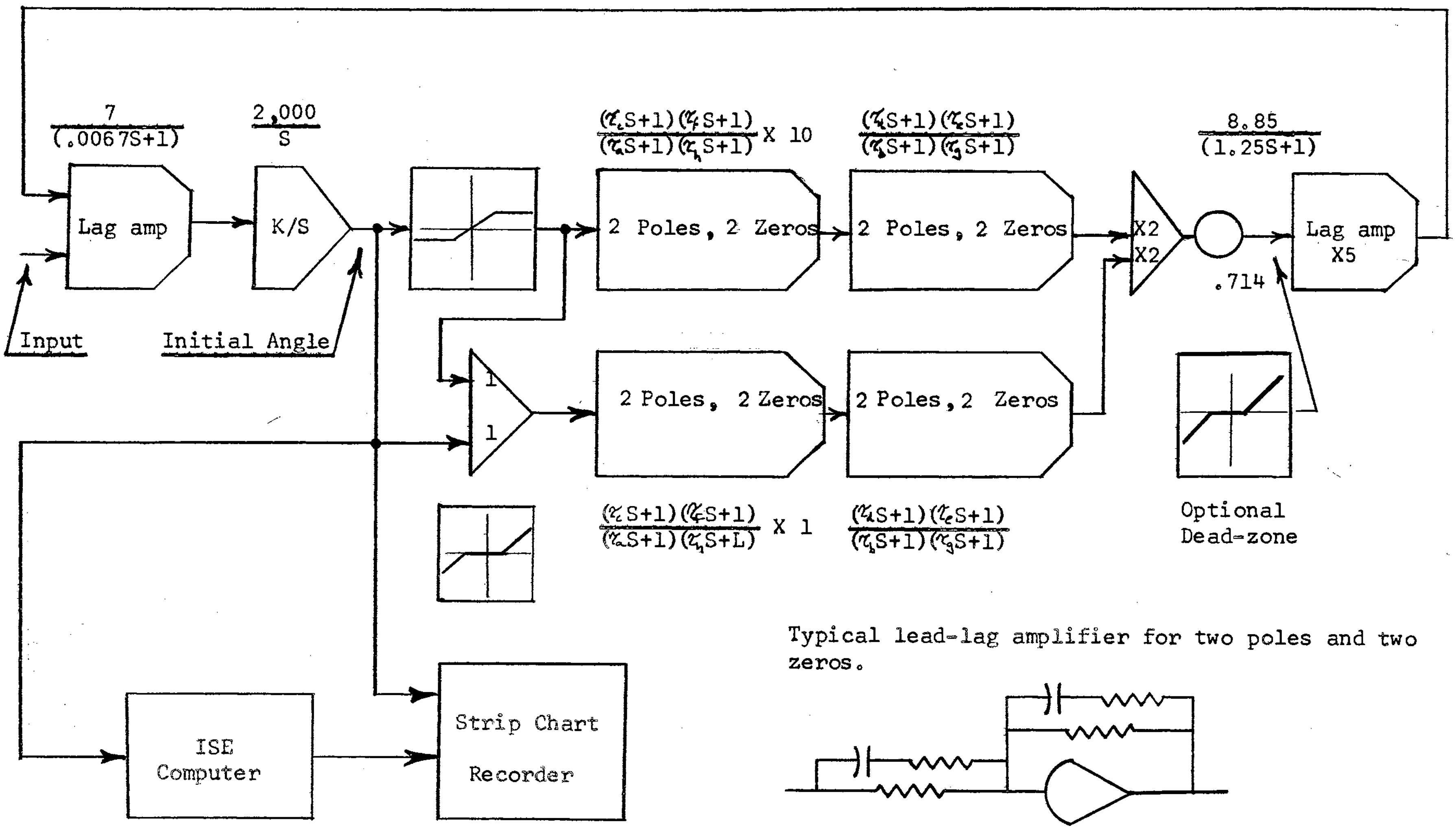

Figure 18. Simulation Diagram for Design Study

The simulation shown in Figure 18 also shows the dead-zone introduced at the secondary plant output used for final comparisons on the composite loop. Plant dead-zone was not considered for the first comparisons.

Performance of Dual-Mode vs. Simple Saturating Loop

As may be observed from the data comparing the saturating loop to that of the dual-mode system, the performance deteriated for saturation levels different than that for which the auxiliary loop was designed. This can be attributed to the fact that the relative signal distribution through the two paths is varied while the compensation remains invariant. Comparative responses are shown in Figure 19 for various saturation levels where the auxiliary loop, it must be remembered, is designed for ten per cent saturation. It may be noted also from these curves that the ISE is reduced as the saturation level approaches that for which the composite loop was designed. Since ten per cent saturation of Loop A alone caused instability, the eleven per cent level was chosen for comparison.

A simple dead-zone representation for kinetic friction in the output element of a control loop leaves much to be desired as an accurate description of the friction phenomena (Rabinowicz, 12). A simple dead-zone was introduced, nevertheless, just prior to the output element transfer function in the analog simulation. The motivation for this decision lay in the idea that there exists at least a first order similarity in the effects upon a loop containing either the simple dead-zone or the more complex description of friction. Thus, if the dual-mode loop configuration will produce adverse effects in a loop with kinetic friction, it should have a similar effect on the loop with a simple dead-zone.

Two noteworthy observations were made concerning the behavior of the

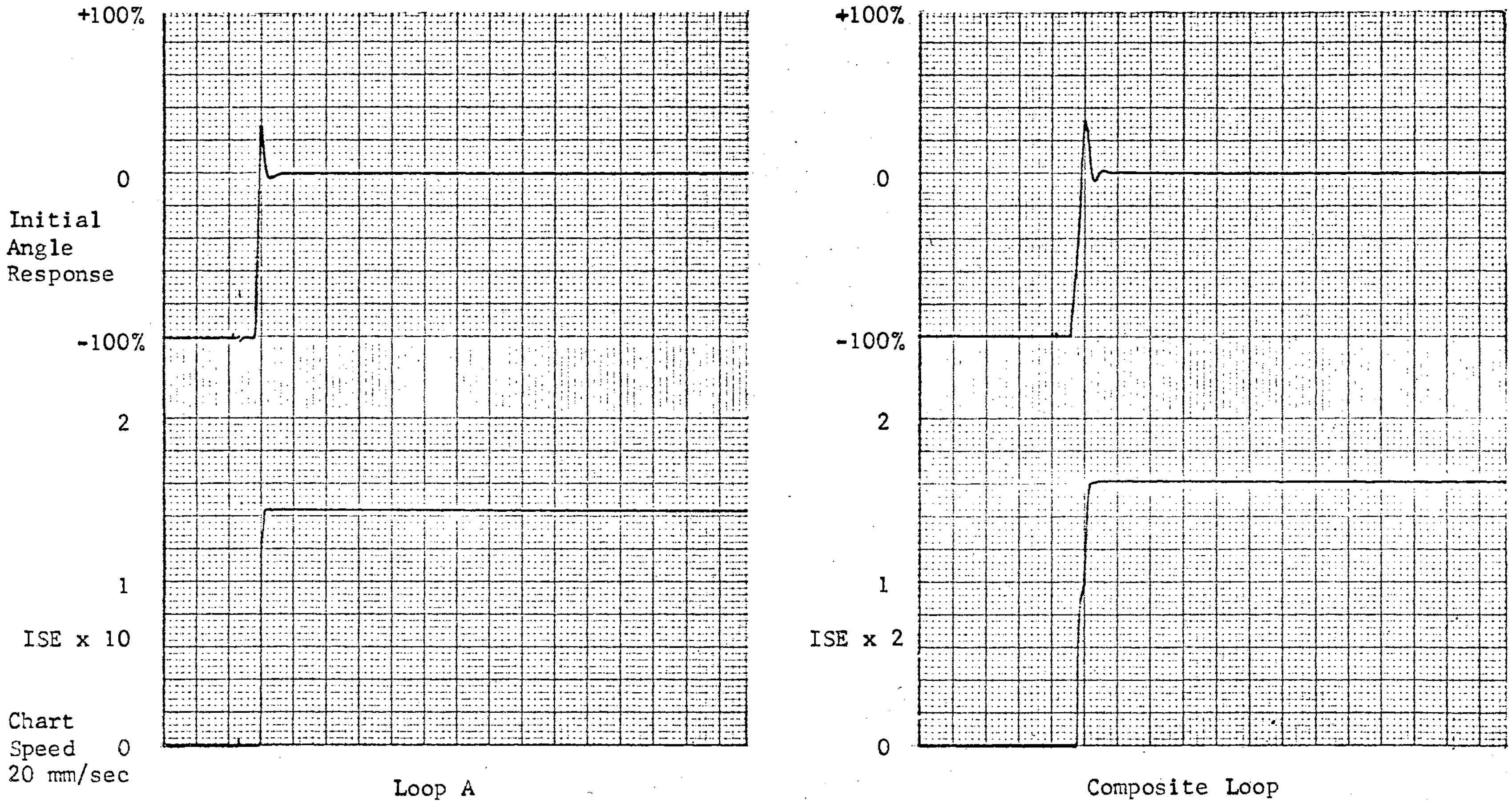

Figure 19a. Comparative Response and Integral of Squared Error for Loop A and Composite Loop for 30% Saturation

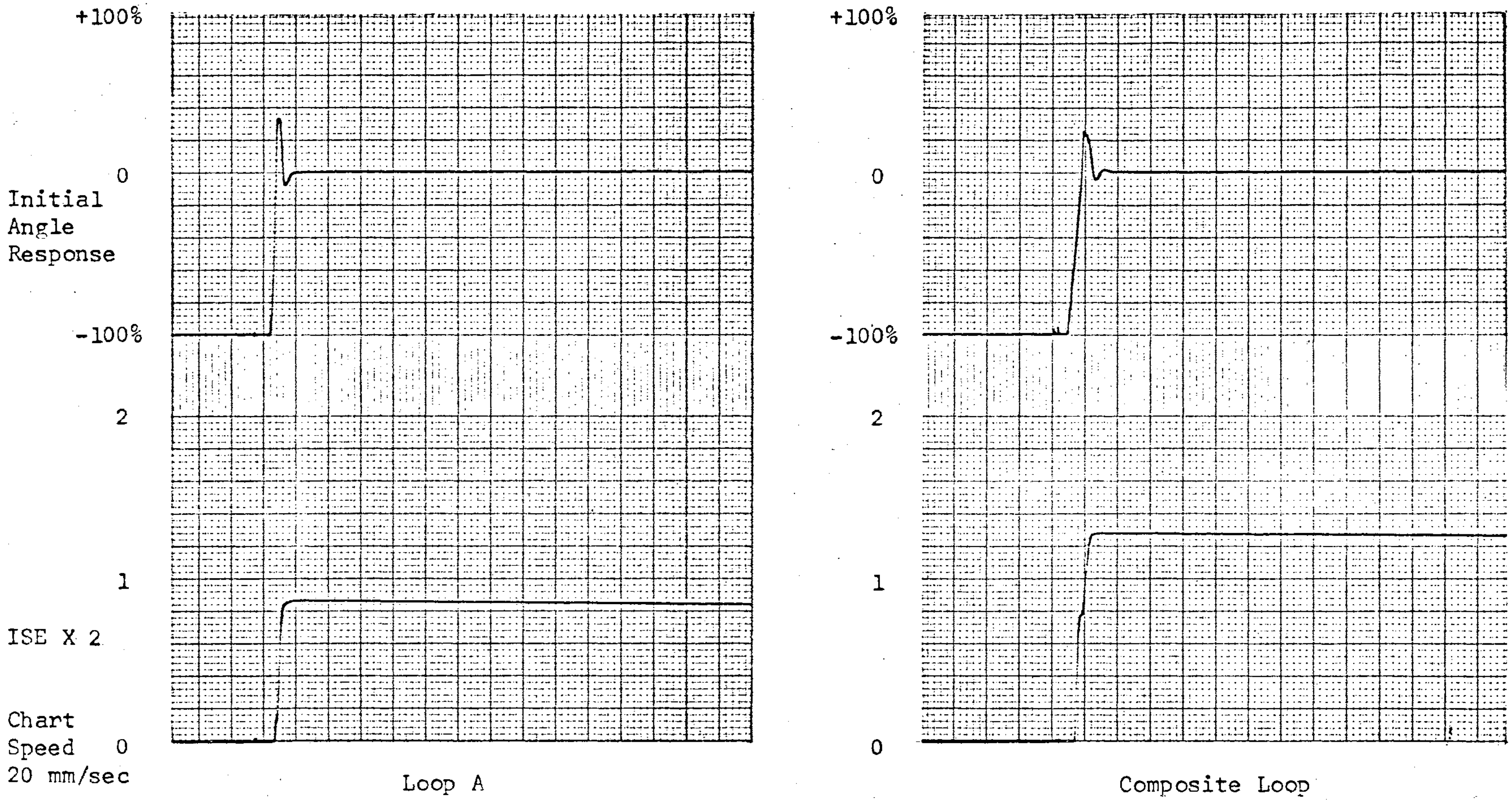

Figure 19b. Comparative Response and Integral of Squared Error for Loop A and Composite Loop for 20% Saturation

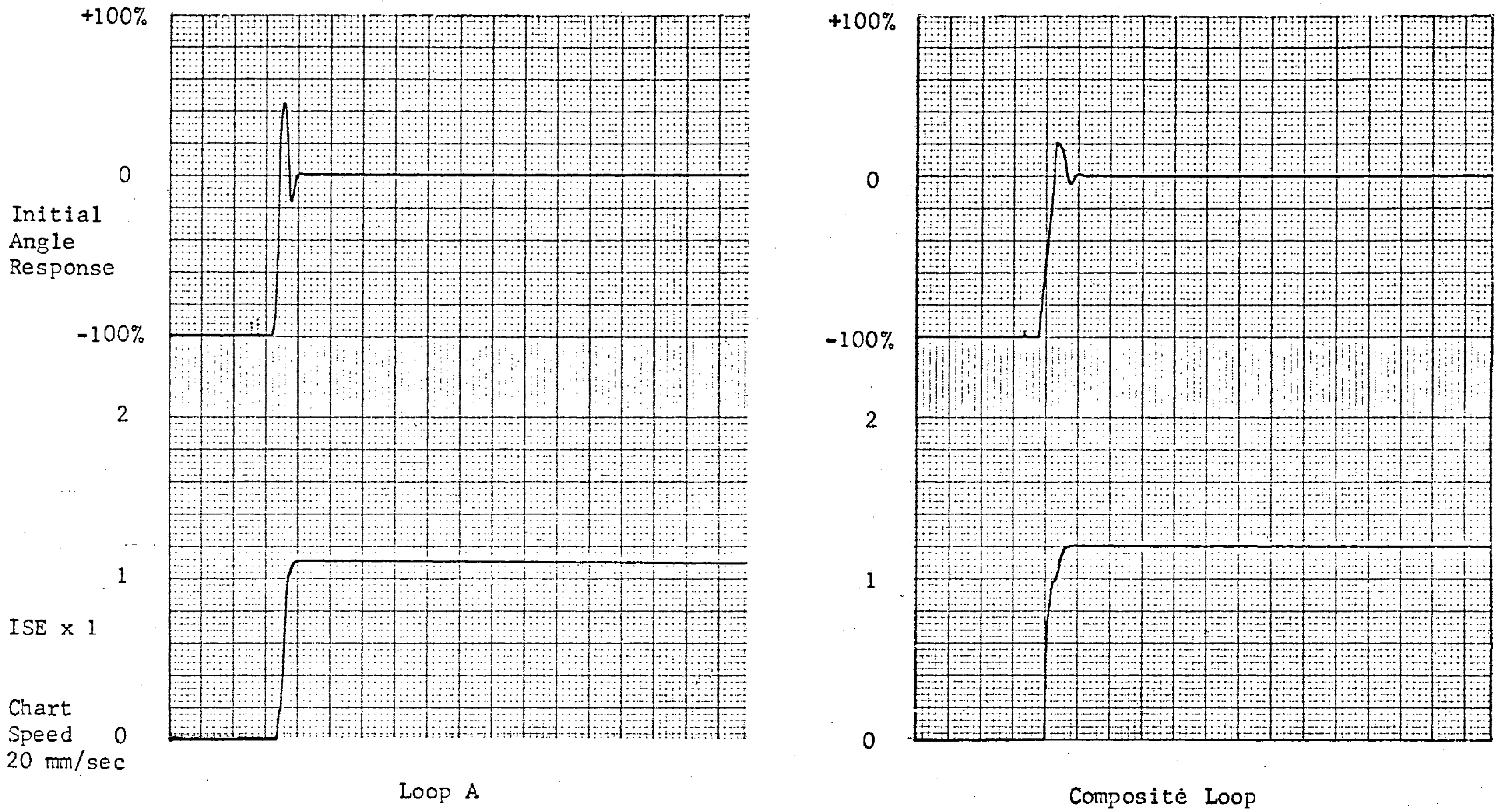

Figure 19c. Comparative Response and Integral of Squared Error for Loop A and Composite Loop for 15% Saturation

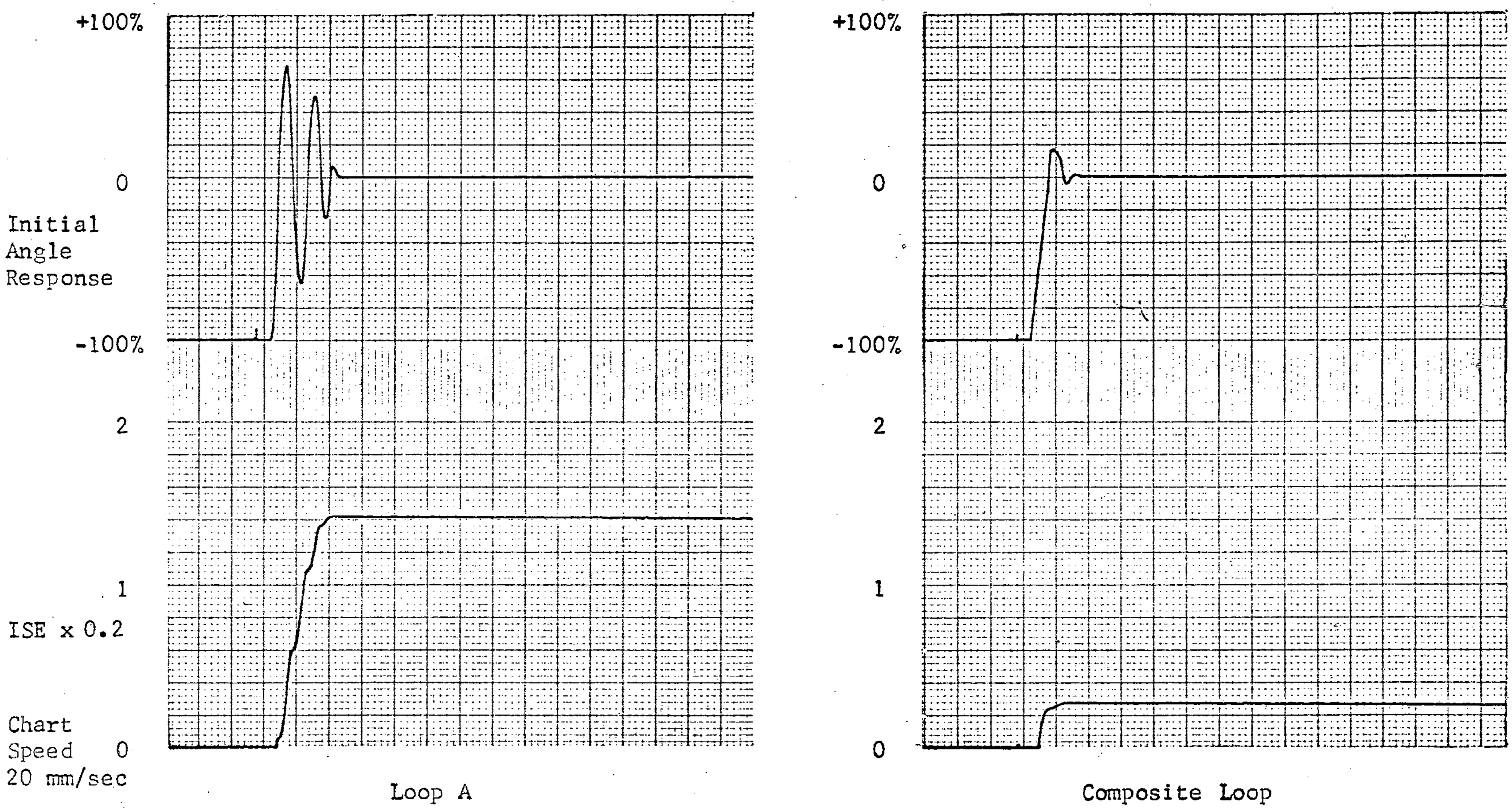

Figure 19d. Comparative Response and Integral of Squared Error for Loop A and Composite Loop for 11% Saturation

composite loop with various dead-zones. Introduction of any amount of dead-zone into the original system (Loop A) raised the level of saturation at which divergence occurred. This was as expected considering the describing function theory described previously. Also, for large values of dead-zone (about five per cent) the effect of the composite loop was the reduction of the limit cycle frequency with little effect on the amplitude of the limit cycle. Meaningful comparisons were difficult to make in this case but two examples are shown in Figure 20.

The most significant result of the test conducted on the dual-mode loop with dead-zone was when the dead-zone level was reduced to a more reasonable two and one half per cent. In this case, as clearly shown in Figure 20, Loop A was not divergent for saturation as little as fifteen per cent but a sustained limit cycle of about four per cent peak value was observed. This behavior was considered to be quite reasonable. When the auxiliary loop was added the dual-mode loop responded with an expected overshoot of about twenty per cent, followed by a complete quenching of the limit cycle, as shown in Figure 20. Since the limit cycle magnitude of Loop A was less than the threshold of the auxiliary loop, it was puzzling to observe an improvement of performance as significant as this. The conclusion was drawn that unbalance in the computer resistors caused the dead-zone or initial slope to be slightly positive instead of zero. An attempt to substantiate this conclusion was made by interchanging the input resistors to the amplifier generating the dead-zone characteristic of Loop B. By interchanging resistors the initial slope should have been made slightly negative and thereby tending to increase the limit cycling instead of quenching it. This theory was abandoned when it was found that the dual-mode loop quenched

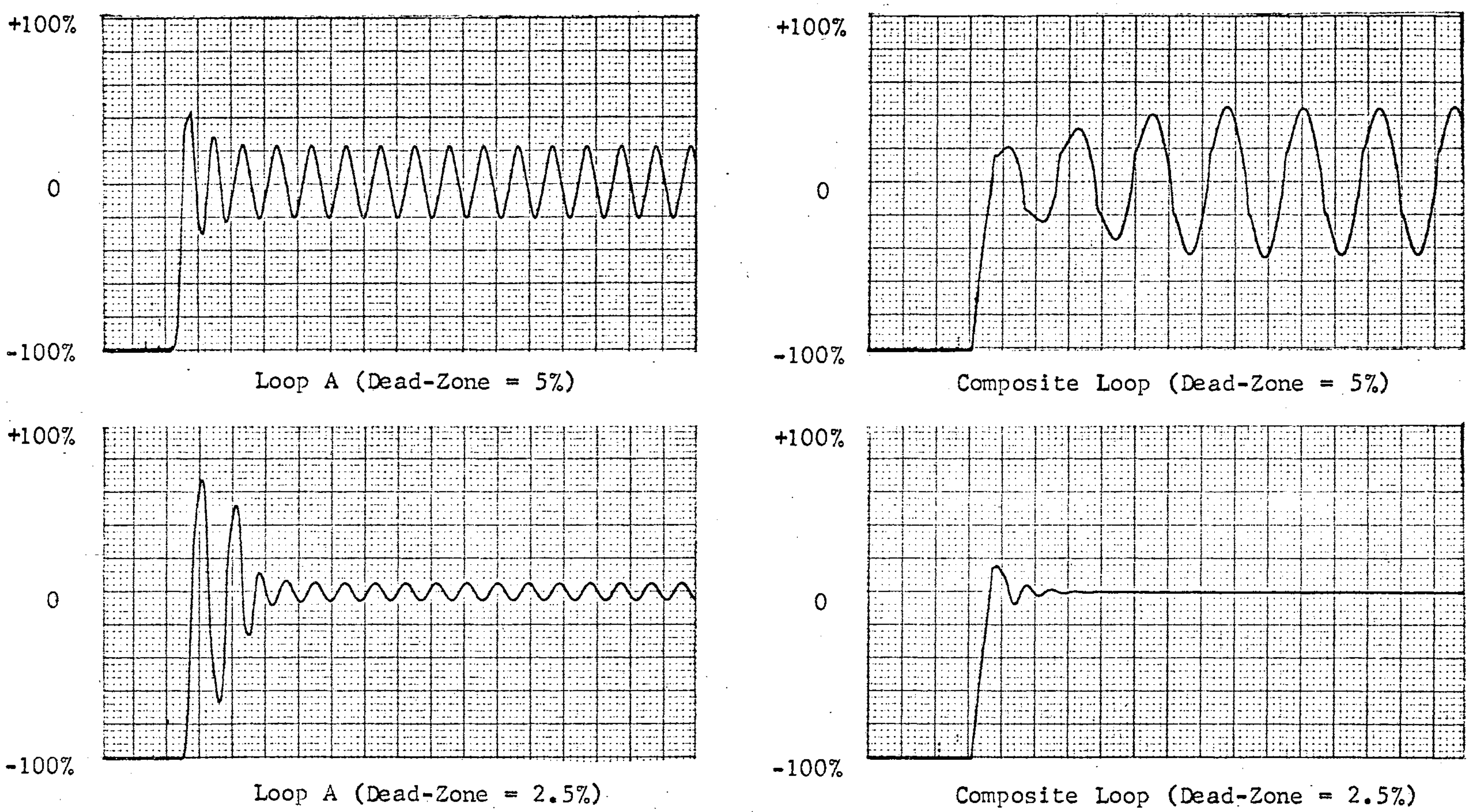

Figure 20. Limit Cycles for Two Different Dead-Zones

the limit cycle for various combinations of resistors. At the writing of this discussion no simple explanation can be offered for this rather significant contribution to the system's performance.

The dual-mode loop seems to have solved many problems in meeting the self imposed constraints of the design problem. Another consideration remains however, and that is maximum error excursion for inputs of various frequencies. Although the performance requirements specify maximum rate inputs only up to one cycle per second, it seemed that a frequency sweep from one tenth cycle to over ten cycles per second at the maximum rated peak amplitude input specified would render a significant comparison. For this comparison, magnitudes of error versus frequency for the dual-mode loop are compared to the error versus frequency for the linear loop (Loop A without saturation). Although both systems exhibited large error excursions for this extreme test, only the composite loop responded without error excursions beyond the destructive (100 per cent initial angle) limit.

The real significance of the dual-mode solution seems to lie in the fact that the performance was achieved without increasing the gain or dynamic voltage range allowable. Neither does this solution assume any additional degree of freedom over the original design. Most adaptive control approaches, according to Horowitz (6), seem to neglect one or both of these limitations.

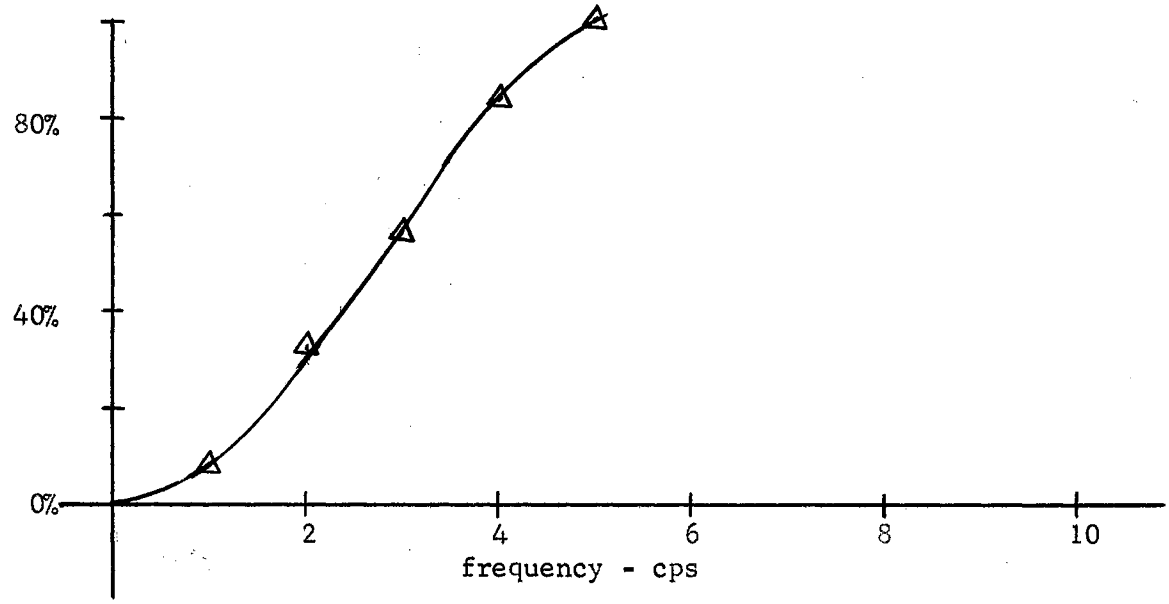

Figure 21a. Peak Error Angle vs. Frequency for Loop A for 0.5 sin(2πft) Input

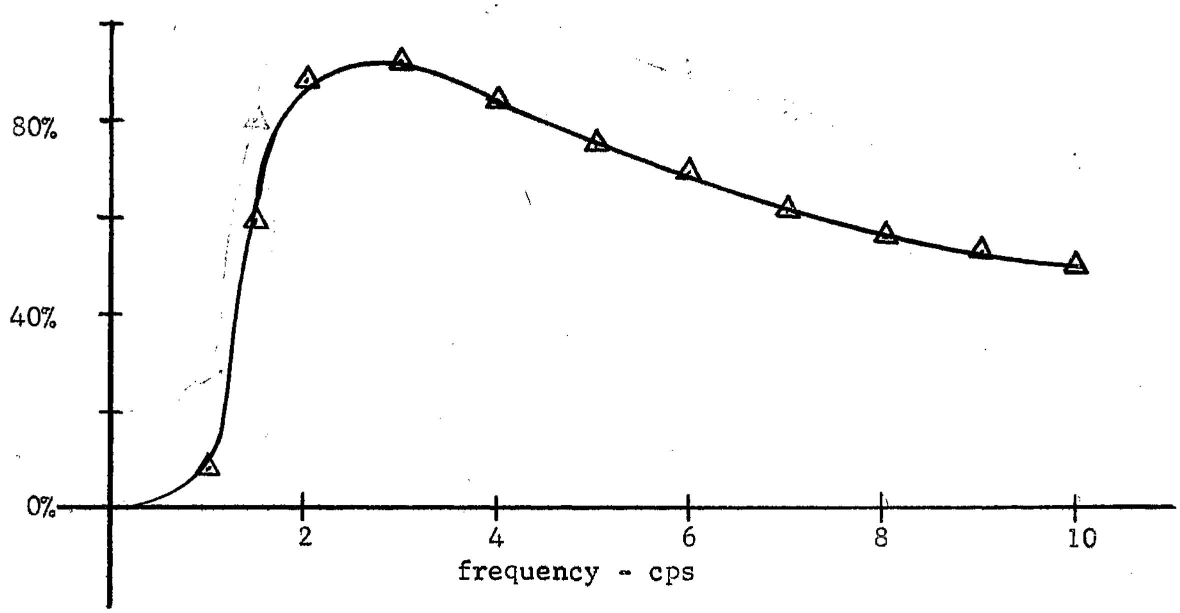

Figure 21b. Peak Error Angle vs. Frequency for Composite Loop for 0.5 sin(2πft) Input

CHAPTER V

SUMMARY AND CONCLUSIONS

A design process involving only limited complexity is presented for the particular problem of eliminating destructive instability in a high gain, narrow band-width, conditionally stable control system. Proper implementation of such a design could conceivably increase the reliability of a control system in that a parallel path of signal transmission is provided.

The subsequent design meets the constraints of limited forward transmission, and invariance of degree of freedom so often ignored by adaptive control designs. Comparisons of the integral squared error (ISE) for the design product indicate that a minimum is approached as a result of judicious selection of parameters.

Finite ISE for unpredictable input signals suggests that the system designed by such a process might have preference over one which became destructively unstable for an unpredicted input.

The proposed approach to a design problem introduced no increase of gain, no differentiation, and no adverse effects on the small signal performance of a loop designed from linear considerations only.

Analog computer simulation studies were quite informative and yielded some very conclusive results. No particular reference was made to problems of simulation peculiar to loops with very high gains. The problems, such as power supply ripple and external noise injection, were noticeably present in the first analog simulation. Since, however, the same problems face the designer of a high gain loop, little need be said about methods of overcoming these problems in the simulation. Judicious placement of components in addition to reduction of the number of amplifiers effected a final simulation of a satisfactory nature. The only significant noise noted on recordings in this study was that produced by the compute-reset relays of the analog computers. This noise is usually noticeable on the recorder traces as a small doublet spike.

The dual-mode solution to a fairly restrictive problem, may for applicable situations, provide the control system designer with an effective method of uncoupling saturation and threshold constraints. Thus the designer is permitted to deal with threshold and small signal characteristics of a control loop almost independently from the large disturbance and loop closing transient characteristics. It has certainly proved to be a satisfactory solution for the design problem presented and amended by various subtle constraints usually attributed to designer bias or integrity.

The design method presented is not a cure-all for control system problems involving nonlinearities. It is, however, a solution to a specific class of nonlinear problems, and as such may be added to the repertoire of the control system designer along with numerous other particular solutions peculiar to specific nonlinear problems.

It is possible, however, that the approach may further stimulate investigation into this area. Hopefully, it may result in a more analytical method of synthesis and analysis.

The idea of the dual-mode suggests, in addition to the composite root locus idea mentioned previously, an analysis using some modified state variable formulation. One might for instance formulate a state variable transition matrix for each individual loop in addition to a weighting matrix to be premultiplied into each transition matrix. Then possibly each stage of transition could be determined by the sum of the two weighted transition matrices. Such a formulation might involve considerable research but might provide a useful analysis tool if accomplished.

SELECTED BIBLIOGRAPHY

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

2. Truxal, J. G. Automatic Feedback Control System Synthesis. New York: McGraw-Hill Book Company, Inc., 1955.

3. Evans, W. R. Control System Dynamics. New York: McGraw-Hill Book Company, Inc., 1954.

4. Harris, F. W. "Passive-Adaptive Compensation Using a Model Reference." M. S. Thesis, Oklahoma State University, (1962), 5.

5. Truxal, J. G. Adaptive Control Systems. Edited by E. Mishkin and L. Braun, Jr. New York: McGraw-Hill Book Company, Inc., 1961, p. 4.

6. Horowitz, I. M. "Plant Adaptive Systems vs. Ordinary Feedback Systems." IRE Transactions on Automatic Control, No. PGAC-7 (January, 1962), 48-56.

7. Lang, G., and J. M. Ham. "Conditional Feedback Systems--A New Approach to Feedback Control." AIEE Transactions Part II, Applications and Industry, Volume 74 (July, 1955), 152-161.

8. Bellman, R. E., and S. E. Dreyfus. Applied Dynamic Programming. New Jersey: Princeton University Press, 1962, pp. 262-263.

9. Aseltine, J. A., A. R. Mancini, and C. W. Sarture. "A Survey of Adaptive Control Systems." IRE Transactions on Automatic Control, No. PGAC-6, (December, 1958), 102-108.

10. Gibson, J. E. Nonlinear Automatic Control. New York: McGraw-Hill Book Company, Inc., 1963, p. 546.

11. Mowrer, O. H. Learning Theory and Behavior. New York: John Wiley and Sons, Inc., 1960, pp. 24-25.

12. Rabinowicz, E. "Stick and Slip." Scientific American, (May, 1956), 109-118.

APPENDIX A

TYPICAL USAGE OF THE GYROSCOPE IN A CONTROL LOOP

Since about 1750, Euler and others have pondered the spinning top, rolling hoop, and similar devices. Not until the World Wars did the gyroscope become important as a tool or system component. Today it is unusual to find an aircraft that doesn't contain at least one gyroscope device. More usually, a host of gyroscopic devices compose a major portion of guidance and stabilization equipment aboard all but the smallest aircraft. Since the end of World War II, marked advances in applicable technology and analysis techniques have led to the development of sophisticated devices beyond the comprehension of the average layman. Such devices comprise the heart of such systems as the guidance for the recently developed atomic powered submarines.

A popular type of gyroscopic device is the single degree of freedom, floated, rate integrating gyro. The rotor of such a device is usually held in a sealed gimbal or "can" and floated in a viscous fluid and confined by a set of pivots and jewels similar to those in a fine watch. Such a device is often specified as the primary component in such devices as stable platforms and accelerometers. Their use in single axis stable platforms, three axis stable platforms, and accelerometers may be cited as the primary source for the motivating design problem of this thesis.

In the case of a single axis platform, the rotations of a body or a set of base coordinates may be isolated from the sensitive axis of a platform by locating on the platform, an integrating rate gyro (IRG). If an electrical pickoff is located on the output axis orthogonal to the spin axis and input axis, any deviation of the gimbal will be sensed, amplified, and fed to the platform driving motor in a manner such that the precession torque on the gyro will balance any torques applied through the platform axis bearings when the base system is rotated. This servo system sets up the equivalent of a much larger gyro wheel capable of rotating the platform merely by its own precession torque. The effect of a servo controlled platform is then to reduce total weight by using very small gyros, and further to introduce only very slight drift effects of the small gyros.

Another popular inertial instrument using the floated gyro is the pendulous gyro accelerometer (PGA) or otherwise called the pendulous integrating gyro accelerometer (PIGA). The acceleration sensing properties come from the addition of an unbalanced mass somewhere along the spin axis of the gyro. When the system is accelerated (or an unbalanced thrust applied) the unbalanced mass tends to swing as a pendulum. The pickoff on the output axis (OA) senses the deviation, excites the drive motor, and sufficient precession torque is developed to exactly cancel the pendulous torque. A diagram of such a device is shown in Figure 22.

If the right hand convention on rotational quantities is used, the spin, input, and output axes respectively form a right-handed co-ordinate system. Since the output axis of the gyro is restrained only by viscous damping, the output angle represents the integral of the net output axis torque on the gimbal. By the electro-mechanical rate rebalance scheme shown in the Figure, the device becomes fundamentally a torque balancing instrument. Thus the analysis of accelerometer action may be pursued.

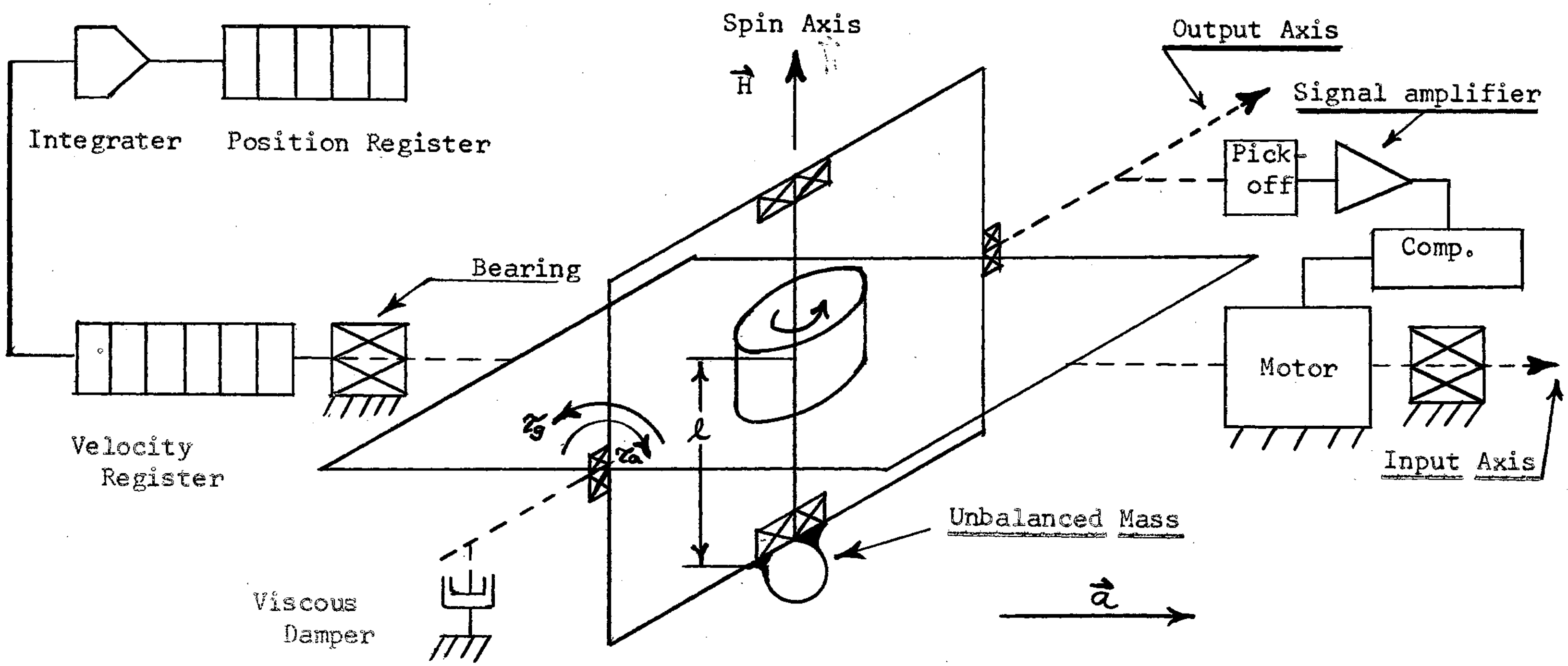

Figure 22. Simplified Schematic of a Pendulous Gyro Accelerometer

First of all, consider the system accelerated in the positive IA direction. The unbalanced mass (m) tends to swing back creating a torque on the output axis equal to the product of mass, length and acceleration. Any rotational displacement about the output axis activates the motor drive along the input axis and develops in the gyro the torque

$$\vec{T}_g = \vec{H} \times \vec{\omega}$$

where H is the vector angular momentum and is equal to the product of the scaler I (rotor moment of inertia) and the vector $\vec{\Omega}$ (gyro spin rate). Thus, since the configuration shown is a nulling device, the effect is to equalize the two torques

$$-\vec{T}_g = \vec{T}_a$$

Thus we may equate the constituent products yielding

$$\vec{H} \times \vec{\omega} = -ml\vec{a} = -p\vec{a}$$

where p = ml is referred to as the pendulosity. Now solving for a, we get

$$\vec{a} = \frac{\vec{H} \times \vec{\omega}}{-p} = -\frac{\vec{H}}{p} \times \vec{\omega}$$

Now the turntable rate may be instrumented to yield acceleration, the turntable angle yields the accumulated velocity, and the integral of the turntable position is the accumulated position. The term H/p is sometimes referred to as the accelerometer scale factor.

To maintain a linear acceleration sensing device, the pendulum must remain perpendicular to the sensing axis. This typically requires very high loop gains between the OA pickoff and the platform drive motor. This in turn implies compensation for stabilizing the closed loop behavior. If, in an acceleration sensing device, only those accelerations which might be expected from thrust sources are to be considered and vibrations eliminated as much as possible, the narrow band-width necessity becomes clear.

APPENDIX B

TRANSFER FUNCTION FOR A DIRECT DRIVE D.C. TORQUE MOTOR

When present state of the art direct drive d.c. torque motors are used with voltage drivers, armature inductance and viscous friction are usually negligible. Armature and driving source resistance, however directly affect the response time. At low motor speeds there is virtually no windage to provide viscous restraint, and armature inductance has been reduced by spacing fewer windings more advatageously.

First, a development for the linear case is considered, then the case where kinetic friction is approximated by a constant retarding torque is developed.

Linear Case

The developed torque is

$$T_d = K_t \, i_a$$

where K_t is the torque constant or sensitivity.

$$I_a = \frac{e_i - e_b}{R_a}$$

neglecting armature inductance.

where

$$e_b = K_v \, p\Theta_o$$

and R_a represents armature plus driving source resistance.

Then

$$i_a = \frac{e_i - K_v\, p\Theta_o}{R_a}$$

The torque delivered to load neglecting viscous restraint and assuming no springing of the shaft is

$$T_1 = J\, p^2\Theta_o$$

Equating torques yields

$$J\, p^2\Theta_o = K_t\, \frac{e_i - K_v p\Theta_o}{R_a}$$

or

$$J\, p^2\Theta_o + \frac{K_t\, K_v\, p\Theta_o}{R_a} = \frac{K_t\, e_i}{R_a}$$

Factoring and collecting yields

$$\frac{p\Theta_o}{e_i} = \frac{K_t}{R_a}\; \frac{1}{(pJ + K_t\, K_v/R_a)}$$

By taking the Laplace transform, the usual form of the transfer function results. That is,

$$\frac{\omega_o}{e_i} = \frac{1/K_v}{\dfrac{J\, R_a}{K_t\, K_v} S + 1}$$

For the motor considered to be the secondary plant of the design study, approximate values were as follows:

$$J = 0.0453 \text{ oz. in. sec}^2$$

$$R = 50 \text{ ohms}$$

$$K_t = 16 \text{ oz. in./ampere}$$

$$K_v = 0.113 \text{ volts/radian/sec}$$

Numerically then:

$$\frac{\omega_o}{e_i} = \frac{8.85}{(1.25\ S + 1)}$$

The development above shows that the response time of the d.c. torque motor and load (secondary plant) is directly dependent upon the four factors: motor load inertia, armature and driving source resistance, torque sensitivity, and the back-emf or generator constant of the motor.

Constant Kinetic Friction Considered

The following development depicts more closely the typical transfer characteristic of a motor-load combination. It shows that constant kinetic friction does not affect the original time function if the friction level is referred back to the voltage input and represented as a dead-zone.

Equating torques as before:

$$J\ p^2\theta_o \pm f_k = K_t/R_a\ (e_i - K_v\ p\theta_o)$$

which yields:

$$\omega_o \left[pJ + \frac{K_t\ K_v}{R_a}\right] = \frac{K_t}{R_a}\, e_i \pm f_k$$

or

$$\frac{\omega_o}{e_i \pm \frac{f_k R_a}{K_t}} = \frac{1/K_v}{\frac{J R_a}{K_t K_v} S + 1}$$

which is valid only for $f_k R_a/K_t$ less than e_i.

From this development, it can be seen how the kinectic friction may be referred back to the voltage input. Also it is shown how certain factors may emphasize or de-emphasize the kinetic friction effect.

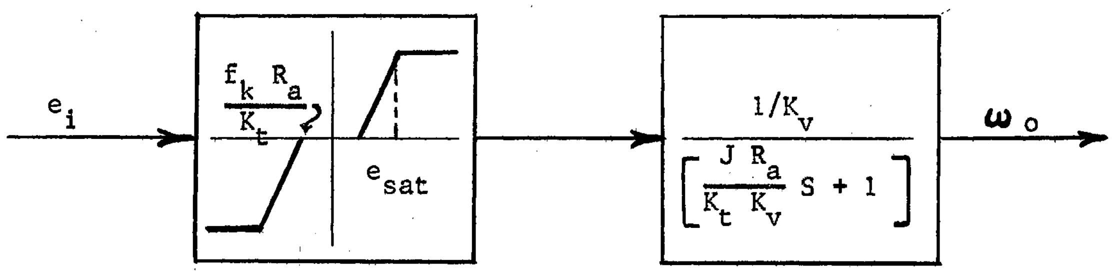

Figure 23. Block Diagram Showing Motor-Load Transfer Function and Lumped Nonlinearities

VITA

Paul Norman Howell

Candidate for the Degree of

Master of Science

Thesis: ADAPTATION AND SATURATION, A DESIGN STUDY

Major Field: Electrical Engineering

Biographical:

Personal Data: Born near Chickasha, Oklahoma, February 17, 1936, the son of Shirley Milton and Faye R. Howell.

Education: Attended grade school in Norman and Chickasha, Oklahoma; attended high school in Okmulgee and was graduated from Chandler High School, Chandler, Oklahoma in May, 1954; received the degree of Bachelor of Science in Electrical Engineering from Oklahoma State University in May, 1959; attended University of Wisconsin part time from September, 1959 to January, 1960; attended University of Minnesota, Extension Division, part time one quarter in 1960; completed requirements for the Master of Science degree at Oklahoma State University in May, 1964.

Professional Experience: Employed by Civil Aeronautics Administration one summer (1958) as student trainee, electronics engineer, in the Facilities Flight Inspection Section; employed by the AC Spark Plug Division, General Motors Corporation, Milwaukee, Wisconsin, from February, 1959 to March, 1960 as a systems engineer; employed by Minneapolis-Honeywell Regulator Company, Aeronautical Division, Minneapolis Minnesota, from April 1960 to September, 1962 as an evaluation engineer, evaluation lead engineer, and design engineer; employed one semester (1964) by Oklahoma State University as a graduate assistant in the controls and circuits laboratories.

Professional Organizations: Member, Institute of Electrical and Electronic Engineers; Engineer-in-Training No. 656, State of Oklahoma.