NASA CONTRACTOR REPORT

NASA CR-823

NASA CR-823

ANALYSIS AND DESIGN OF SPACE VEHICLE FLIGHT CONTROL SYSTEMS

VOLUME IV - NONLINEAR SYSTEMS

by Arthur L. Greensite

Prepared by
GENERAL DYNAMICS CORPORATION
San Diego, Calif.
for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • JULY 1967

NASA CR-823

ANALYSIS AND DESIGN OF SPACE VEHICLE FLIGHT CONTROL SYSTEMS

VOLUME IV - NONLINEAR SYSTEMS

By Arthur L. Greensite

Issued by Originator as GDC-DDE65-056

Prepared under Contract No. NAS 8-11494 by
GENERAL DYNAMICS CONVAIR
A DIVISION OF GENERAL DYNAMICS CORPORATION
San Diego, Calif.

for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

FOREWORD

This report was prepared under NASA Contract NAS 8-11494 and is one of a series intended to illustrate methods used for the design and analysis of space vehicle flight control systems. Below is a complete list of the reports in the series:

Volume I	Short Period Dynamics
Volume II	Trajectory Equations
Volume III	Linear Systems
Volume IV	Nonlinear Systems
Volume V	Sensitivity Theory
Volume VI	Stochastic Effects
Volume VII	Attitude Control During Launch
Volume VIII	Rendezvous and Docking
Volume IX	Optimization Methods
Volume X	Man in the Loop
Volume XI	Component Dynamics
Volume XII	Attitude Control in Space
Volume XIII	Adaptive Control
Volume XIV	Load Relief
Volume XV	Elastic Body Equations
Volume XVI	Abort

The work was conducted under the direction of Clyde D. Baker, Billy G. Davis and Fred W. Swift, Aero-Astro Dynamics Laboratory, George C. Marshall Space Flight Center. The General Dynamics Convair program was conducted under the direction of Arthur L. Greensite.

TABLE OF CONTENTS

TABLE OF CONTENTS, Contd

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS, Contd

LIST OF ILLUSTRATIONS, Contd

1. INTRODUCTION

This monograph discusses a variety of techniques available for the analysis of nonlinear control systems. While there is voluminous technical literature on this subject, it is, for the most part, of a theoretical nature beset with serious computational difficulties in application.

It is the primary purpose of this volume to consider those techniques which have proved most useful in the analysis of launch vehicle control systems. Well established techniques such as phase plane and describing function will be treated from first principles together with extensive application to problems of flight control. Various refinements described in the recent literature will be treated.

A substantial portion of the monograph deals with the Lyapunov theory of stability with emphasis on flight control applications. While this technique is still under active development, its broad features are well established, and it has proved useful in providing insight to previously obscure areas.

The general theme of this volume is that of displaying the features of nonlinear flight control problems obtainable by judicious application of modern nonlinear control theory, indicating both the power and potential of these methods together with a discussion of limitations and areas for future research.

2. STATE OF THE ART

Nonlinear control theory is a subject of active research at the present time. It is safe to say that this particular branch of control theory will never reach the definitive form which characterizes linear control. This is of course due to the fact that a nonlinear system is best defined in a negative sense; that is, a nonlinear system is merely one that is not linear. As a result, any effort whose purpose is to yield useful results, must sharply restrict the scope of the study. The most general theory available is that of Lyapunov. The analytical elegance of this concept is accompanied by serious difficulties in application due mainly to the broad generality of the theory; in short, its chief virtue is also its fundamental handicap. Much progress has, however, been obtained by delineating specialized cases which of course yield specialized results. The best example in this area is the Lur'e technique. There is no doubt that further results of theoretical and practical interest will be forthcoming in the future.

Of the specialized methods presently available for pragmatic design, the most prominent are the phase plane and describing functions. Recent developments in the former have been concerned with improving the methods for constructing the phase plane portraits and extending their application. The generality of the describing function technique has been extended by the introduction of the "Dual Input Describing Function."

The fundamental restriction in the phase plane method is the limitation to second order systems, while the use of describing functions involves the assumption of low band pass properties of the system being analyzed.

However, the formulation of any mathematical model involves some compromise with reality. Ultimately, the experience and judgment of the control engineer provide the best guide for the proper application of these methods together with their proper interpretation.

The principal value in the proper application of nonlinear techniques is the generation of qualitative features of a control system which are essentially unaffected by inclusion of higher order dynamic effects and which thereby permit a rational interpretation of results obtained by computer simulation of a complex system.

3. RECOMMENDED PROCEDURES

3.1 PHASE PLANE ANALYSIS

The phase plane method is concerned with determining the general features of the solution of the differential equation

$$\ddot{x} + f(x, \dot{x})\,\dot{x} + g(x, \dot{x}) = 0 \tag{1}$$

The phase plane is a plot of x versus $\dot{x}$ and represents the trajectory of the system for any given set of initial conditions, x (o) and $\dot{x}$ (o). By classifying and studying the general form of these trajectories, one may draw conclusions on the system behavior without having to solve the equation for every set of initial conditions. While these methods are both elegant and powerful, their application to the analysis of nonlinear flight control problems is severely limited by three restrictions:

1. The phase plane is applicable only to second order systems.
2. Only systems that are not externally excited may be studied.
3. The range of admissable nonlinearities is limited.

Each of these limitations can to some extent be circumvented. Extensions of the method to a phase space of order higher than two have been developed by Ku,[1] Grayson and Mishkin[2] and Kuba and Kazda,[3] among others. There is however a substantial increase in complexity both at the computational level and in physical insight. As a result, these methods have not been used extensively in practical problems.

The second limitation may be partially removed to the extent that step and ramp inputs may be treated. To see this, let Eq. (1) be written as

$$\ddot{x} + f(x, \dot{x})\,\dot{x} + g(x, \dot{x}) = a + bt \tag{2}$$

where a and b are constants and $t > 0$.

Making the change of variable, $z = x - a - bt$, Eq. (2) becomes

$$\ddot{z} + f(\dot{z} + b, z + a + bt)(\dot{z} + b) + z = 0 \tag{3}$$

This equation is autonomous* if either $b = 0$ (step function alone) or $f(x, \dot{x}) \equiv f(\dot{x})$.

*i.e., t does not appear explicitly.

The third limitation is inherent in the form of the differential equation, (1); namely, only signal-dependent nonlinearities are allowed. The coefficients of x and $\dot{x}$ may not contain the independent variable explicitly.

In spite of the rather limited range of systems thus amenable to analysis, the phase plane method is valuable in that it yields qualitative results for systems of higher order and is exact for those cases which are adequately described by second order equations. In the latter case, a wealth of insight into phenomena which cannot be predicted even approximately by the linear theory is obtained, and this often provides a framework for predicting and explaining in qualitative fashion those control systems in which are included higher order dynamic effects.

The phase plane method is based on three essential concepts:

1. The phase plane representation.
2. Singular points.
3. Limit cycles.

These will be discussed in the following sections.

3.1.1 Phase Portrait

By defining a new variable, y, Eq. (1) may be written as

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -f(x, y)\, y - g(x, y) \end{aligned} \tag{4}$$

Actually, the method to be discussed is applicable to the slightly more general system described by

$$\begin{aligned} \dot{x} &= P(x, y) \\ \dot{y} &= Q(x, y) \end{aligned} \tag{5}$$

It is easy to eliminate t completely from this set of equations; viz.

$$\frac{dy}{dx} = \frac{Q(x, y)}{P(x, y)} \tag{6}$$

A variety of specialized techniques have been developed for determining, graphically, the phase plane trajectories of systems represented by Eq. (1) or Eq. (6). They differ mainly in the specific form of the equation treated. Since these methods are dealt with at length in standard texts[4], only two of these will be described briefly.

3.1.1.1 Graphical Methods

A variety of graphical methods are available to solve Eq. (6) depending on the specific form of this equation. Perhaps the simplest technique is the so-called "Isocline Method" which is applicable if $P(x, y) = y$ in Eq. (6). This method is perhaps best explained using a specific example. Consider the equation*

$$\ddot{x} - 0.2 (1 - x^2) \dot{x} + x = 0 \tag{7}$$

This may be written in the form

$$\frac{dy}{dx} = \frac{0.2 (1 - x^2) y - x}{y} \tag{8}$$

where $y = \dot{x}$. A series of lines are drawn for different constant values of $\frac{dy}{dx} \equiv m$. These are shown in Fig. 1 for $m = 0, \pm 0.5, \pm 1$. Now let A be a point on the solution curve. (This may represent the initial state $x(o), y(o)$.) The slope at point A is equal to 1 since it lies on the curve for $m = 1$. Point B on the system trajectory may be determined as follows. We note first of all that the trajectory proceeds in a clockwise direction, as may be readily verified by considering the relation between x and y. Now since B is on the $m = 0.5$ curve, its slope equals 0.5, and the average slope between A and B is therefore 0.75. Drawing a line of slope = 0.75 through point A consequently determines point B. Proceeding in this fashion, the entire trajectory is formed as shown in Fig. 1.

The accuracy of the method can be made as high as desired by using a large number of isoclines (lines of constant slope).

Another method which is somewhat more direct proceeds by constructing the system trajectory by means of small circular arcs. This approach, known as the "Delta Method," is applicable to equations of the form

$$\ddot{x} + \varphi(\dot{x}, x) = 0 \tag{9}$$

By properly defining a variable δ, we may write this as

$$\ddot{x} + \delta + x = 0 \tag{10}$$

---

*The knowledgeable reader will recognize this as the van der Pol equation, the workhorse of nonlinear theory.

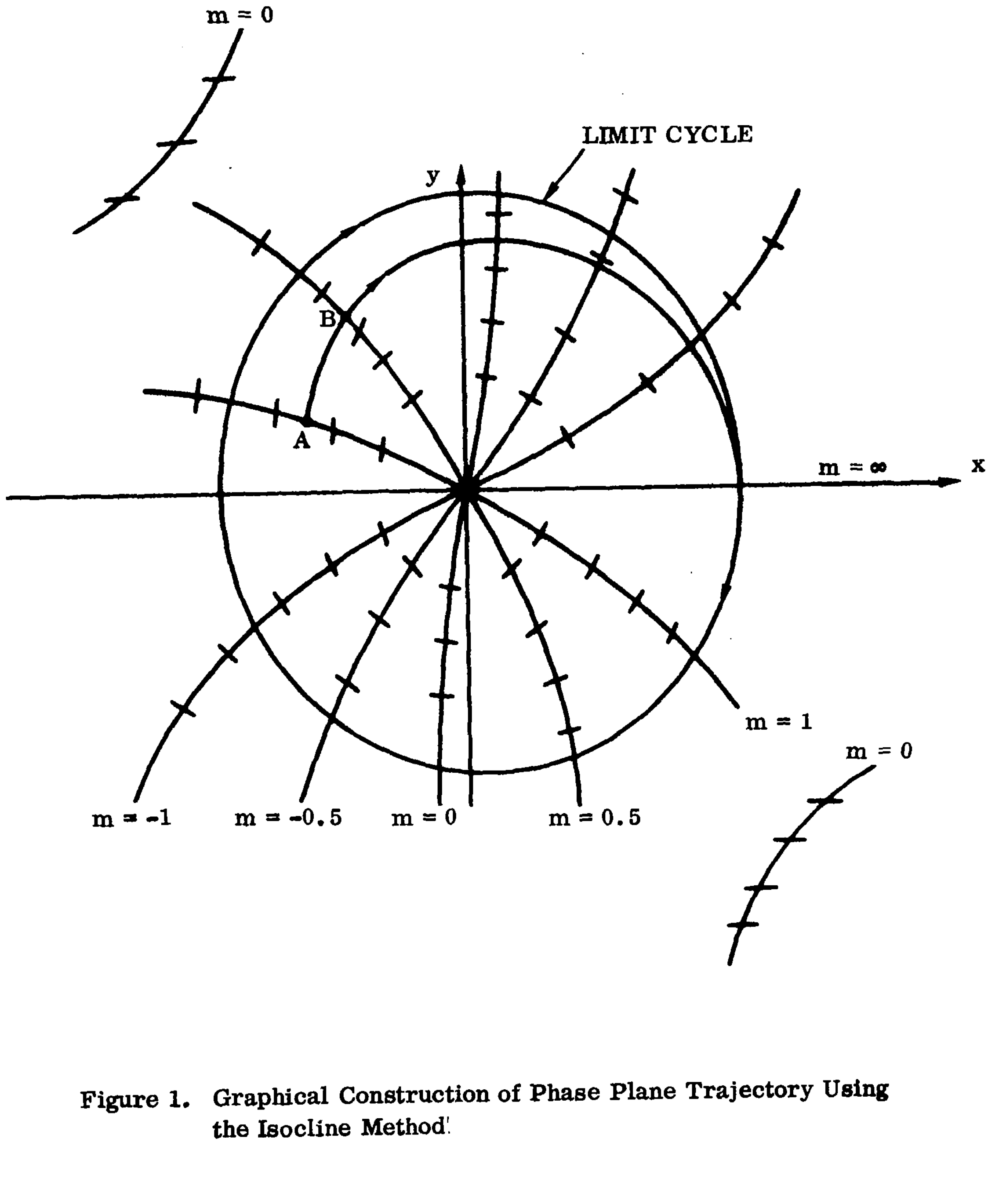


Figure 1. Graphical Construction of Phase Plane Trajectory Using the Isocline Method.

Again using $y = \dot{x}$, we may write Eq. (10) as

$$\frac{dy}{dx} = -\frac{x + \delta}{y} \tag{11}$$

where

$$\delta = \varphi(x, y) - x \tag{12}$$

For small changes in the variables x and y, δ remains essentially constant and in this case, Eq. (11) may be integrated to yield

$$y^2 + (x + \delta)^2 = R^2 \tag{13}$$

where R^2 is a constant. This is the equation of a circle of radius R whose center is at $x = -\delta$, $y = 0$. Thus, for a suitably small increment, the system trajectory is a circular arc of the form shown by Eq. (13). The construction proceeds as follows. The initial point is located at $x(o)$, $y(o)$, as shown in Fig. 2. These values of x and y determine the value of δ which thereby locates the center of the circular arc and determines the radius R automatically. A short circular arc is then drawn to locate the new point, x_1 and y_1. The procedure is then repeated until the entire trajectory is formed. Fig. 3 illustrates the trajectory obtained for the system

$$\ddot{x} + 25\,(1 + 0.1\,x^2)\,x = 0$$

with the initial conditions

$$x = -1.8$$

$$\dot{x} = y = -1.6$$

A typical circular arc construction is shown for the point $x = 3$, $y = 0$ for which the value of δ equals 2.7.

Remark: Graphical methods, by themselves, are of limited usefulness in establishing the essential properties of the nonlinear equation as represented by Eq. (5) or Eq. (6). Such questions as stability, boundedness of solutions, etc. are determined in rigorous and elegant fashion by invoking the concepts of singular points and limit cycles which are discussed in Sections 3.1.2 and 3.1.3. A powerful technique for piecewise linear systems is treated in Section 3.1.4. As a preliminary to this, the phase plane solutions for second order linear systems are developed in the next section.

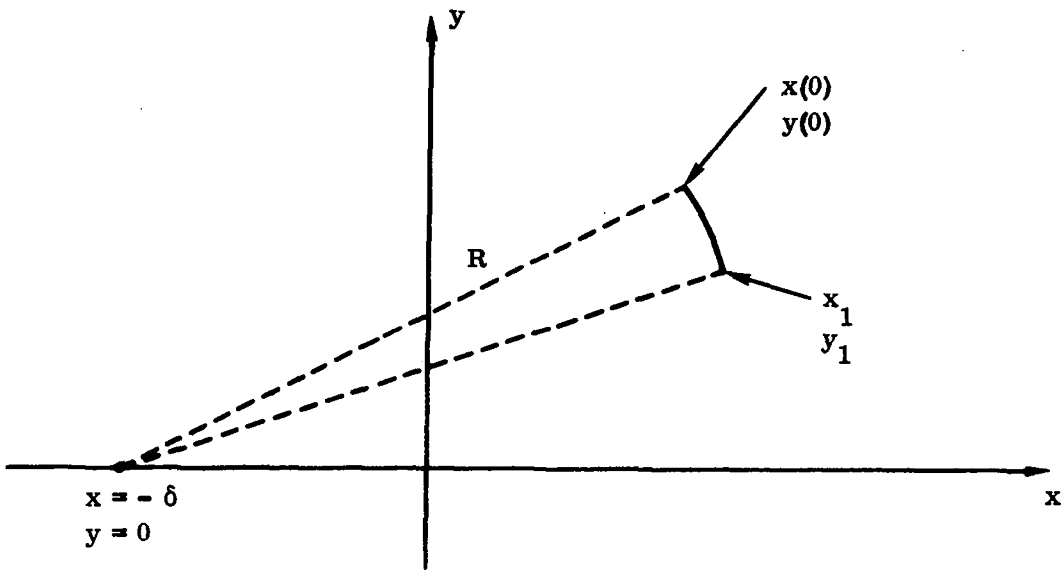


Figure 2. Basic Construction of the Delta Method

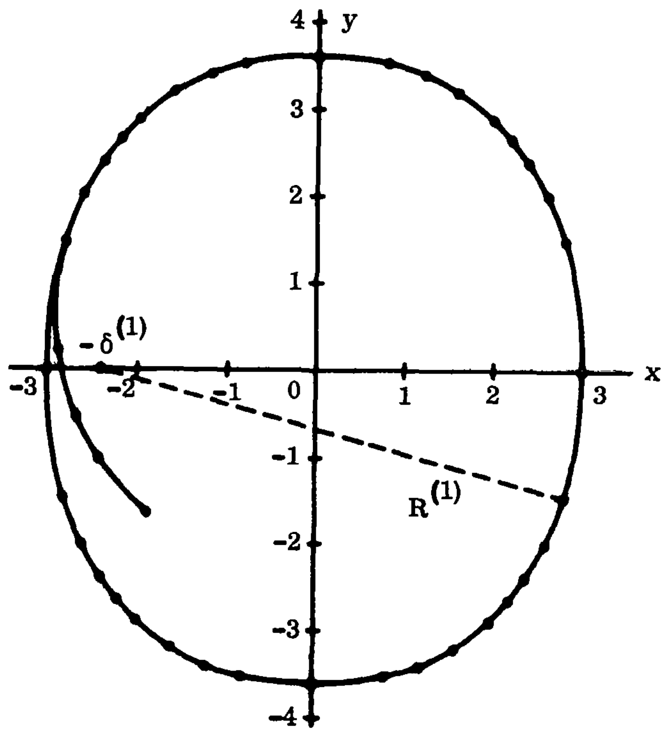


Figure 3. Complete Trajectory Constructed Using the Delta Method

3.1.1.2 Direct Solution of Equations

In certain cases, the trajectory of the system in the phase plane may be obtained by integrating Eq. (6) directly. For example, consider the equation

$$\ddot{\epsilon} + 2\zeta\omega_o\dot{\epsilon} + \omega_o^2\epsilon = \omega_o^2\epsilon_c \tag{14}$$

$$\begin{aligned} \epsilon_c &= A \qquad t \geq 0 \\ &= 0 \qquad t < 0 \end{aligned}$$

which represents a linear system with a step input.

Writing

$$x = \epsilon - A$$

$$\tau = \omega_o t$$

this may be put in the form

$$x'' + 2\zeta x' + x = 0 \tag{15}$$

where primes denote differentiation with respect to τ. Letting $y = x'$ and dividing through by y, Eq. (15) may be expressed as

$$\frac{dy}{dx} = -\frac{2\zeta y + x}{y} \tag{16}$$

If $\zeta = 0$, then this integrates immediately to

$$x^2 + y^2 = c^2 \tag{17}$$

which is the equation of a circle centered at the origin; the constant c is determined by the initial conditions.

If $\zeta = 0$ and the sign in front of the x term in Eq. (15) were minus instead of plus, then (15) would integrate to

$$y^2 - x^2 = c^2 \tag{18}$$

or

$$x^2 - y^2 = c^2 \tag{19}$$

depending on the value of the initial conditions. These curves, which are rectangular hyperbolas, are shown in Figs. 4 and 5.

Turning now to the general case, we consider the equations

$$\begin{aligned} \dot{x} &= y \\ \dot{y} + 2\zeta\omega y + \omega^2 x &= 0 \end{aligned} \tag{20}$$

which may be expressed as

$$\frac{dy}{dx} = -\frac{(2\zeta\omega y + \omega^2 x)}{y} \tag{21}$$

after eliminating t. Performing a change of variable

$$y = vx \tag{22}$$

Eq. (21) becomes

$$\left[v + 2\zeta\omega + \frac{\omega^2}{v}\right] dx = -x\,dv \tag{23}$$

This may now be integrated by separation of variables to yield

$$y^2 + 2\zeta\omega xy + \omega^2 x^2 = C_1 e^{\varphi(x,\,y)} \tag{24}$$

$$\varphi(x,\,y) = \frac{2\zeta}{\sqrt{1-\zeta^2}} \tan^{-1}\left(\frac{y + \zeta\omega x}{\omega x\sqrt{1-\zeta^2}}\right)$$

when $|\zeta| < 1$.

For the case of $|\zeta| > 1$, Eq. (23) integrates to

$$(y - \lambda_1 x)^{\lambda_1} = C_2\,(y - \lambda_2 x)^{\lambda_2} \tag{25}$$

$$\lambda_1 = -\zeta\omega - \omega\sqrt{\zeta^2 - 1}$$

$$\lambda_2 = -\zeta\omega + \omega\sqrt{\zeta^2 - 1}$$

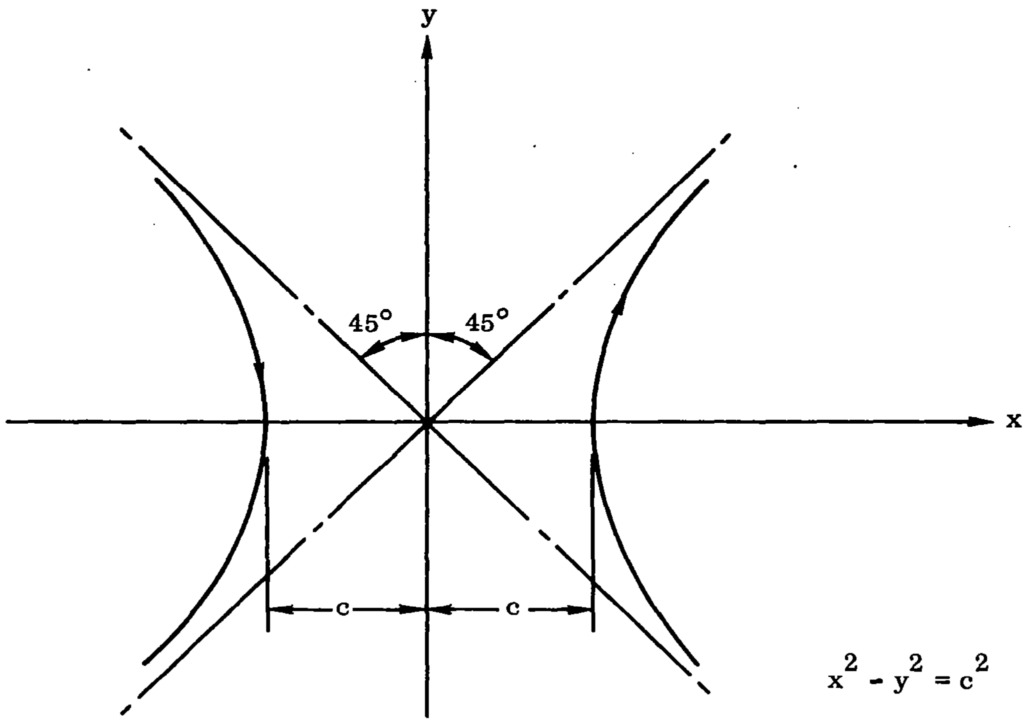


Figure 4. Rectangular Hyperbola of Eq. (18)

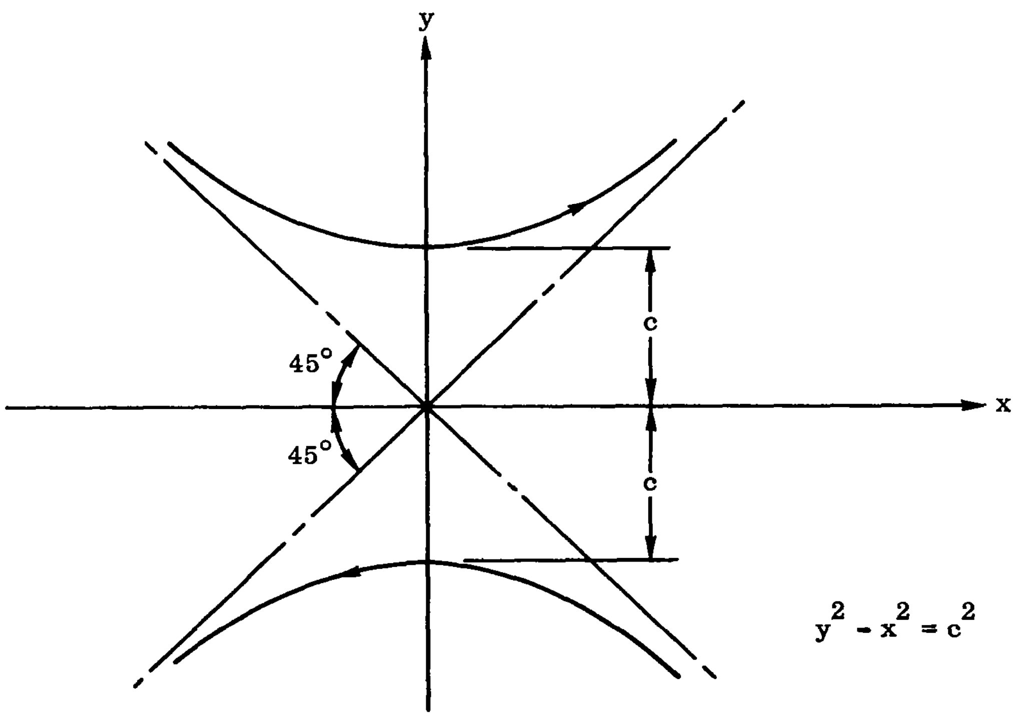


Figure 5. Rectangular Hyperbola of Eq. (19)

Typical trajectories corresponding to Eq. (24), the underdamped case, are shown in Figs. 6 and 7 for $\omega = 1$ and $\zeta = 0.3$ and 0.5, respectively. From Eq. (21), observe that the slope changes sign at $y = -x/2\zeta$. This line is also shown in the figure.

The trajectories corresponding to Eq. (25), the overdamped case, are shown in Figs. 8 and 9 where $\omega = 1$ and $\zeta = 1.4$ and 2.0, respectively. The phase portrait is characterized by principal directions and the final portion of any trajectory is along a straight line through the origin. The principal directions are determined as follows. Suppose that it is possible to choose a constant, λ, such that $y/x = \lambda$ is a portion of the trajectory. For this to be true, we find from Eq. (21) that the following condition must be satisfied

$$\lambda^2 + 2\zeta\omega\lambda + \omega^2 = 0$$

This is in fact the characteristic equation of the system. In other words, when λ is a root of the system characteristic equation, then there is a solution in the phase plane given by $y = \lambda x$. There are two such paths, corresponding to λ_1 and λ_2. These paths are straight lines through the origin and they define principal directions of the phase portrait. Just as the time behavior of the overdamped second order system is primarily determined, initially, by the larger of the two roots, and finally by the smaller one, so a typical trajectory in the phase plane starts parallel to the principal direction with the larger slope (say λ_2) and ends parallel to the direction with the smaller slope (λ_1). These features of the phase portrait for overdamped systems are evident in the trajectories shown in Figs. 8 and 9.

The phase portraits of Eqs. (24) and (25) may be further simplified by an appropriate change of variable. Consider first Eq. (24). If new variables are defined by

$$u = \left(\omega\sqrt{1-\zeta^2}\right)x$$
$$v = y + \zeta\omega x$$

then (24) may be expressed as

$$u^2 + v^2 = C_1 e^{\varphi(u,\, v)}$$

$$\varphi(u,\, v) = \frac{2\zeta}{\sqrt{1-\zeta^2}} \tan^{-1}\left(\frac{v}{u}\right)$$

with a change to polar coordinates

$$u = \rho \cos\theta$$
$$v = \rho \sin\theta$$

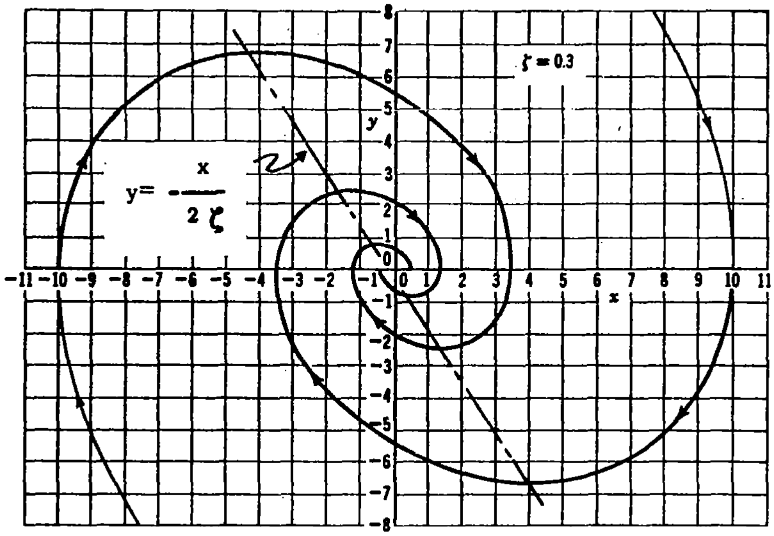


Figure 6. Phase Portrait of the Damped Second-Order System
$\ddot{x} + 2\zeta\dot{x} + x = 0;\ \zeta = 0.3$

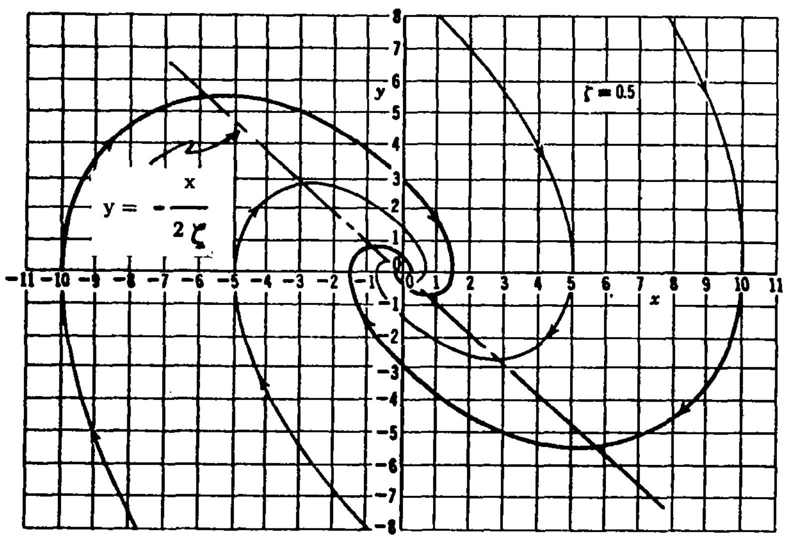
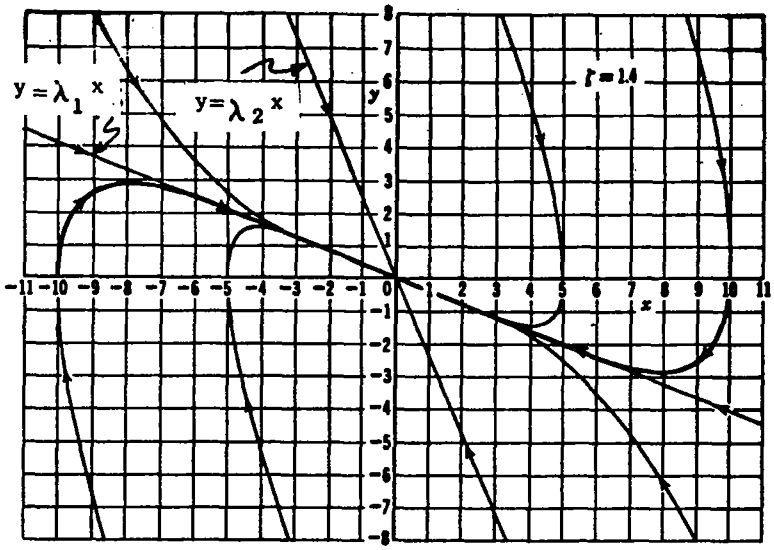
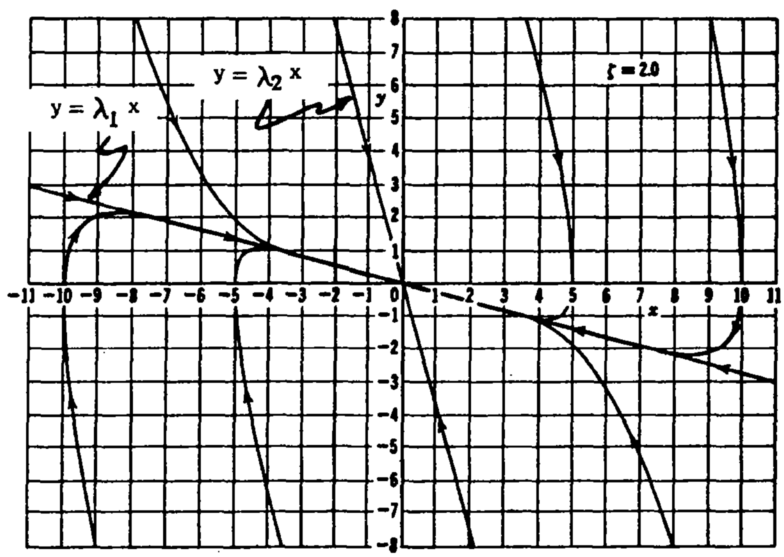


Figure 7. Phase Portrait of the Damped Second-Order System
$\ddot{x} + 2\zeta\dot{x} + x = 0;\ \zeta = 0.5$



$|\lambda_1| < |\lambda_2|$

Figure 8. Phase Portrait of the Damped Second-Order System
$\ddot{x} + 2\zeta\dot{x} + x = 0;\ \zeta = 1.4$



$|\lambda_1| < |\lambda_2|$

Figure 9. Phase Portrait of the Damped Second-Order System
$\ddot{x} + 2\zeta\dot{x} + x = 0;\ \zeta = 2.0$

and this becomes

$$\rho = C_3 e^{\theta}$$

which is the equation of a logarithmic spiral. Turning to Eq. (25), if one adopts the coordinate transformation

$$u = y - \lambda_2 x$$

$$v = y - \lambda_1 x$$

then it is found that Eq. (25) reduces to

$$v = C_4 u^{\lambda_2/\lambda_1}$$

which is the equation of a family of parabolic-like curves centered on the v axis.

Remark: The importance of the results derived in this section lies not in their neatness and elegance but rather in the fact that they are basic to the study of the piecewise linear systems considered in Section 3.1.4. What is significant is that the phase portraits are distinctive, depending upon whether the system is underdamped, overdamped, or undamped. These qualitative characteristics, taken in conjunction with properties related to singular points and limit cycles, provide one of the most valuable tools available for predicting the qualitative features of non-linear phenomena. These topics are taken up next.

3.1.2 Singular Points

The slope given by Eq. (6) becomes indeterminate when simultaneously

$$\begin{aligned} P(x, y) &= 0 \\ Q(x, y) &= 0 \end{aligned} \tag{26}$$

The point at which this occurs is termed a "singular point." An analysis of the nature of singular points is important from the point of view of equilibrium states of the system; whether trajectories converge to or diverge from this state, or in some cases, yielding information on the existence of closed paths in the phase plane (periodic phenomena).

If a singular point exists at $x = a$ and $y = b$, then a Taylor expansion of Eq. (5) about this point yields

$$\dot{x} = a_1 (x - a) + a_2 (y - b) + \text{higher order terms}$$

$$\dot{y} = b_1 (x - a) + b_2 (y - b) + \text{higher order terms}$$

In a sufficiently small region about the singular point, the higher order terms are negligible and the system behaves linearly, viz.

$$\dot{x} = a_1 (x - a) + a_2 (y - b) \tag{27a}$$

$$\dot{y} = b_1 (x - a) + b_2 (y - b) \tag{27b}$$

For this case, the characteristic equation of the system is found to be

$$s^2 - (a_1 + b_2) s + (a_1 b_2 - a_2 b_1) = 0 \tag{28}$$

The nature of the roots of this equation determines the behavior of the system trajectories in the phase plane. Denoting these roots by λ_1 and λ_2, six cases may be distinguished.

I. λ_1 and λ_2 both real and negative. This case is termed a stable node and all trajectories converge to the singular point as shown in Fig. 10a.

II. λ_1 and λ_2 are complex conjugate with negative real part. This condition gives rise to a stable focus with all trajectories converging to the singular point in the manner shown in Fig 10b.

III. λ_1 and λ_2 are pure imaginary. For this case, the motion is simple harmonic with the oscillation amplitude dependent on the initial conditions. The trajectories are closed paths around the singular point which in this case is termed a center. The phase plane portrait is shown in Fig. 10c.

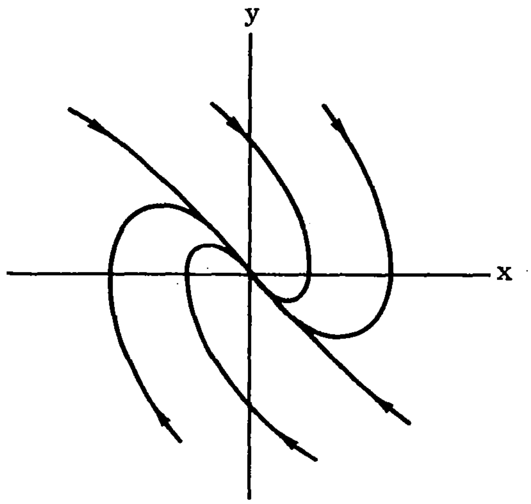
IV. λ_1 and λ_2 both real and positive. This gives rise to an unstable node with all trajectories diverging from the singular point as shown in Fig. 10d.

V. λ_1 and λ_2 are complex conjugate with positive real part. This case results in oscillatory motion but with divergent amplitude; all trajectories leave the singular point as shown in Fig. 10e. The singular point here is termed an unstable focus.

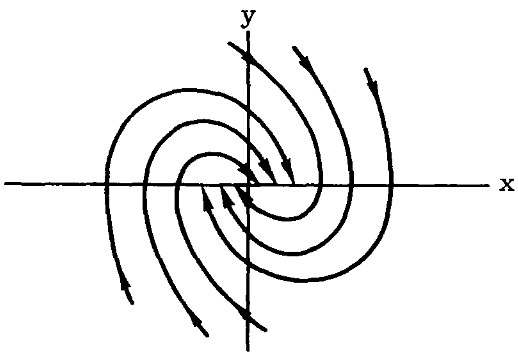
VI. λ_1 and λ_2 both real with one positive and one negative. The singular point here is termed a saddle point. The trajectories are depicted in Fig. 10f.

The type of singular point encountered is essentially dependent on the coefficients in Eq. (28). If we write

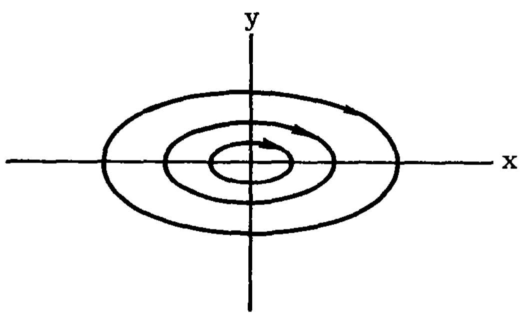
$$\mu = a_1 b_2 - a_2 b_1$$
$$\nu = - (a_1 + b_2)$$



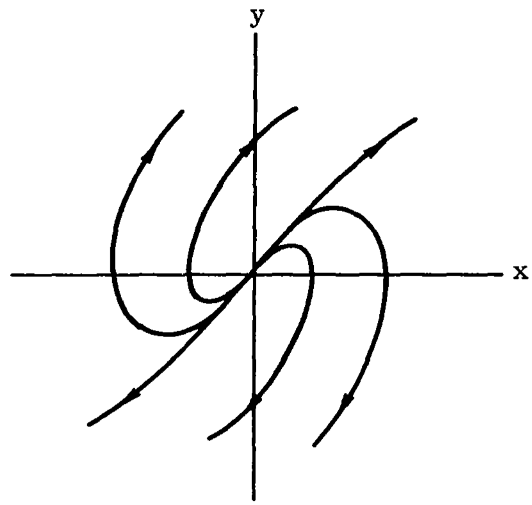
a. Stable Node - The roots are both real and negative



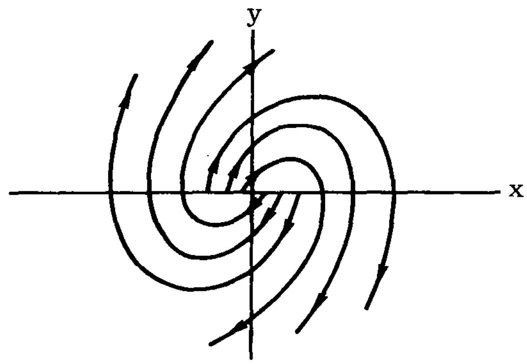
b. Stable Focus - The roots are complex conjugates with negative real parts



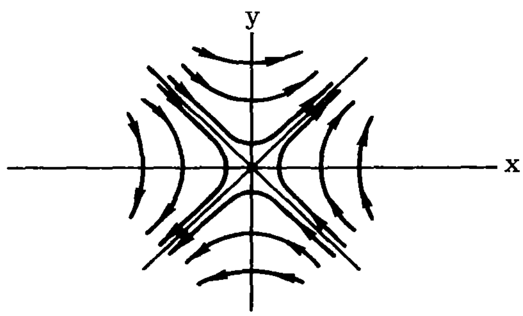
c. Center - The roots are complex conjugates with zero real parts



d. Unstable Node - The roots are both real and positive



e. Unstable Focus - Roots are complex conjugates with positive real parts



f. Saddle Point - The roots are both real and of opposite sign

Figure 10. Typical Phase Plane Trajectories for Singular Points

Then a concise tabulation of the singular points is as depicted in Fig. 11, where the line of demarcation between the nodes and foci is given by the parabola

$$\nu^2 = 4\mu$$

The actual equations of the phase plane trajectories are obtained by substituting Eqs. (27) for P(x, y) and Q(x, y) in Eq. (5) and writing the resulting set in the form of Eq. (6) after eliminating t. Eq. (6) then has a solution in closed form in the manner discussed in Section 3.1.1.2. Note that while the results obtained there were valid for the whole phase plane, the case discussed here is limited to the immediate vicinity of the singular point because of the linearizing process involved in deriving Eqs. (27).

3.1.3 Limit Cycles

The determination of the nature of the singular points together with the use of the graphical methods described previously is sufficient to indicate the general nature of the system trajectories. However, for purposes of stability analysis, a knowledge of the possible existence of limit cycles is essential. A limit cycle is an isolated closed path in the phase plane. The limit cycle may be stable or unstable depending on whether the paths in the neighborhood converge to the limit cycle or diverge away from it. From a practical point of view, only the former is important. The two types of limit cycles are illustrated in Fig. 12.

Unfortunately, there is no completely general method for determining the limit cycles of any given system. The available methods are summarized below in four theorems which are stated without proof. *

Theorem I: If outside a circle C_1 in the phase plane all paths are converging (the radial distance to the point moving along the path is decreasing, the radial distance being measured to the center of the circle, C_1), and inside a smaller circle C_2, with the same center as C_1, the paths are diverging, then a limit cycle must exist within the region bounded by C_1 and C_2.

Theorem II (Bendixson's First Theorem): With P(x, y) and Q(x, y) defined as in Eq. (5), then no limit cycle can exist within any region in which $\left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right)$ does not change sign.

Theorem III (Bendixson's Second Theorem): If a path stays inside a finite region, D, and does not approach a singular point, then it must either be a limit cycle or approach a limit cycle asymptotically.

Theorem IV (Poincare′): Within any limit cycle, the number of node, focus, and center types of singularity must exceed the number of saddle points by one.

---

* The proofs may be found in Ref. 5.

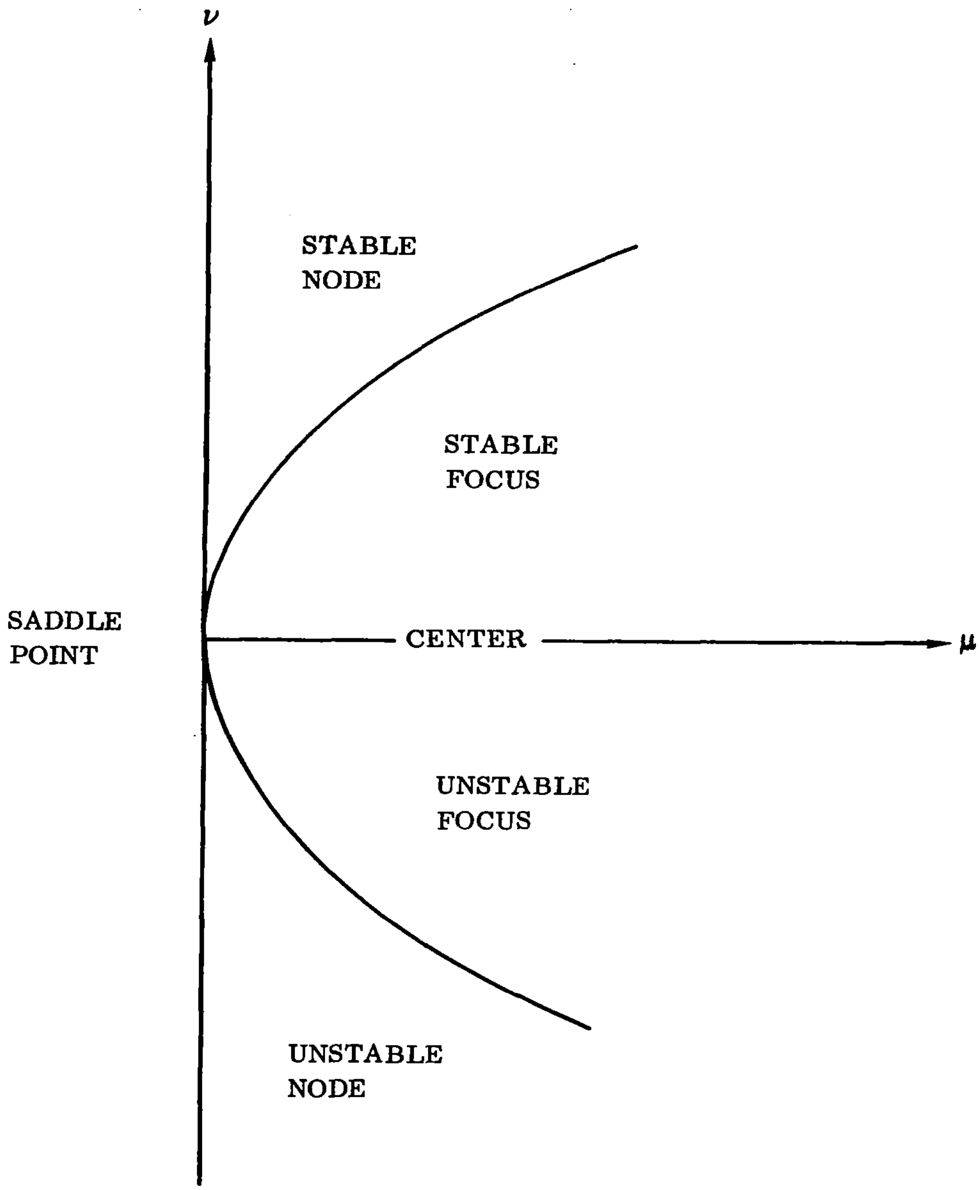


Figure 11 Singular Points as a Function of μ and ν

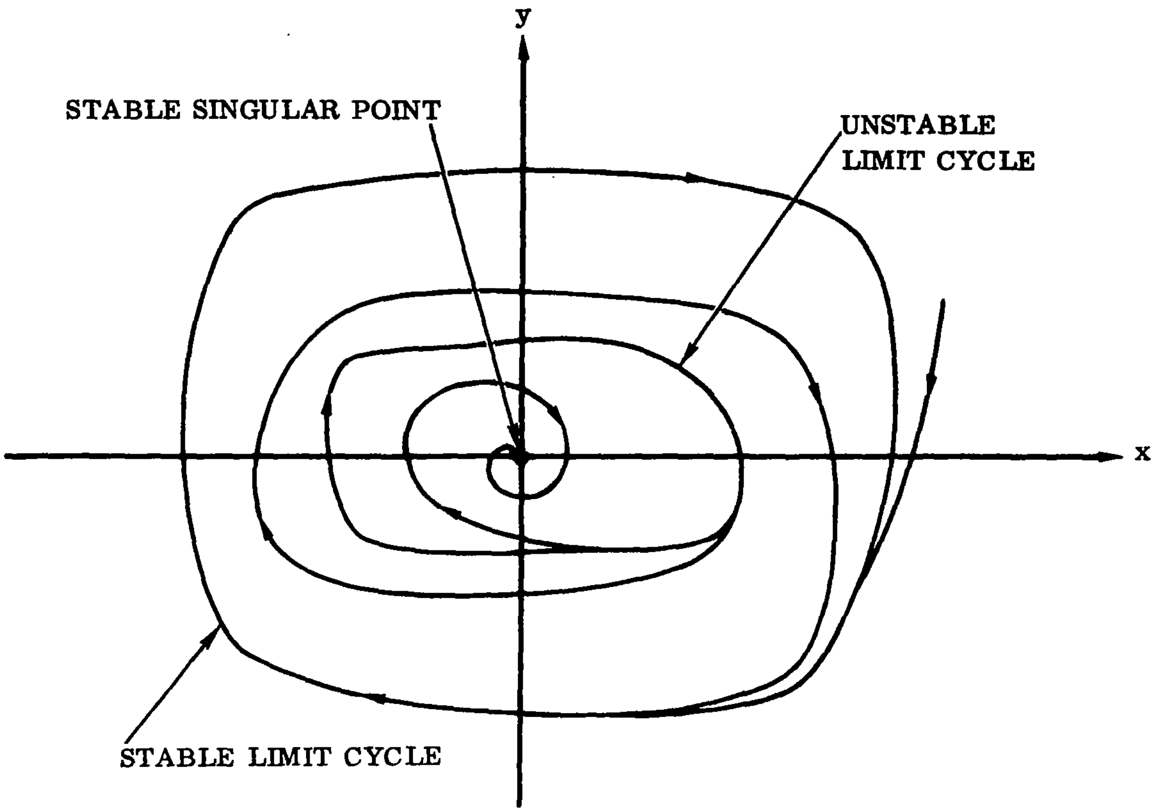


Figure 12. Types of Limit Cycles

Remark: Most systems of interest, which can be put in the form (5), may be analyzed by invoking the concepts of singular points and limit cycles, together with appropriate graphical methods. The form of the phase portrait thus determined yields valuable information relating to stability, points of equilibrium, boundedness or periodicity of the solution, etc. These phenomena often cannot be predicted by linearized techniques alone. The application of these ideas will be illustrated by two simple examples.

Example 1: We consider the van der Pol equation

$$\ddot{x} - \epsilon (1 - x^2) \dot{x} + x = 0$$

Writing this in the form

$$\dot{x} = y \equiv P(x, y)$$
$$\dot{y} = \epsilon (1 - x^2) y - x \equiv Q(x, y)$$

it is readily found that a singular point exists at $x = y = 0$ and that this singular point represents an unstable node. Furthermore

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} = \epsilon\ (1 - x^2)$$

By Theorem II, a limit cycle exists since this quantity changes sign at $x = \pm 1$. The phase portrait for this system is shown in Fig. 13. It is worthy of note that this limit cycle is reached whatever the form of the initial conditions. Linear systems do not display this phenomenon.

Example 2: The system described by

$$\ddot{x} + \omega_o^2 x - h^2 x^3 = 0$$

closely approximates the motion of a pendulum with large deflection angles. Writing this as

$$\dot{x} = y \equiv P(x, y)$$
$$\dot{y} = -(\omega_o^2 - h^2 x^2)\, x \equiv Q(x, y)$$

we see that there are three singular points as follows:

a. A center at $x = y = 0$

b. A saddle at $x = \omega_{o/h}$, $y = 0$

c. A saddle at $x = -\omega_{o/h}$, $y = 0$

This information is sufficient to construct the phase portrait shown in Fig. 14. Note that depending on the initial conditions, there is either a stable limit cycle or a divergent motion. This again is in contrast with linear systems where the stability is not affected by the initial conditions.

Various other cases of phase plane analysis are treated in the references. The work of Davis [6] is especially noteworthy in its analysis of some complex situations where the nonlinearities are analytic and differentiable. Many cases in flight control systems are characterized by piecewise linear conditions; viz.; relays, saturation, dead zone, hysteresis, etc. In other cases, the nonlinearity may be approximated by piecewise linear segments. The analysis of this type of system is particularly simple and elegant, yielding considerable information on the properties of the motion with comparatively little effort. This approach, which is due to Kalman,[7] is described next.

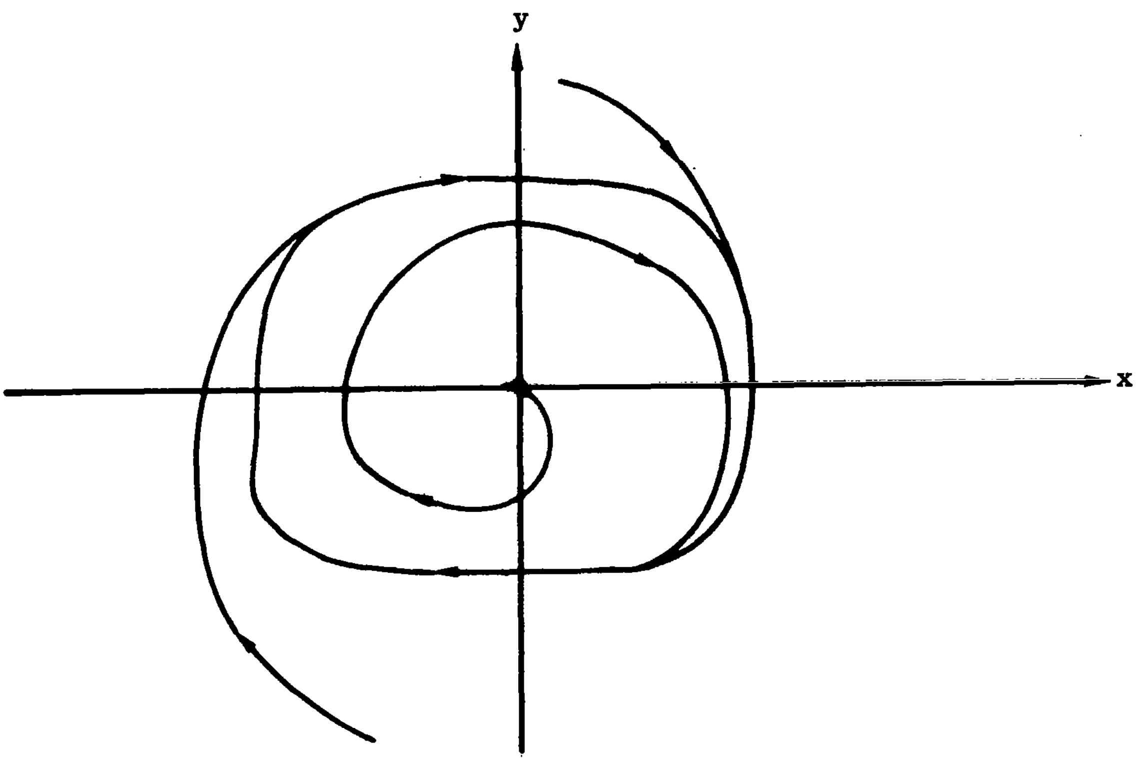


Figure 13. Phase Portrait for Example 1

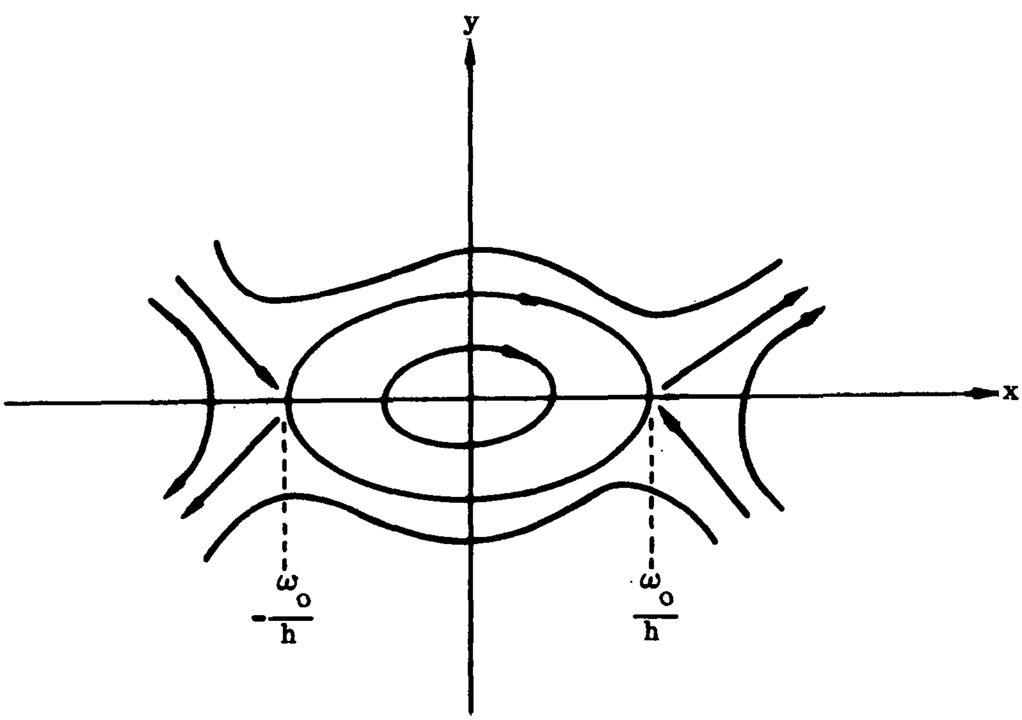


Figure 14. Phase Portrait for Example 2

3.1.4 Piecewise Linear Systems

For a linear system, a determination of the nature of the critical point is sufficient to characterize the complete behavior of the system for all time. This problem was discussed in Section 3.1.1.2. Fig. 10 depicted the types of phase plane trajectories which characterize each type of singular point. In Section 3.1.2, these trajectories were said to indicate the form of the phase portrait in the vicinity of the singular point since the relevant equation was linearized about this point. However, if the system is linear in the entire region, then these paths are valid in the entire region. These considerations suggest that the phase plane be decomposed into regions, in each of which, the system is linear, then combining the various paths in order to completely describe the motion. At all times, the qualitative aspects of the system response will be kept in the foreground. The aim is to obtain the greatest possible degree of insight into the problem with a minimum of labor. While exact numerical results may be obtained with some additional effort, this is not our primary objective at the moment. As will be seen, a wealth of information on properties of the system motion will be derived with little computational effort.

The approach will be described by applying the method to several problems in the field of flight control. A few preliminary remarks are in order. For any region in the phase plane where the system is linear, the motion is governed by

$$\ddot{\theta} + 2\zeta\omega\dot{\theta} + \omega^2\theta = \omega^2\theta^* \tag{29}$$

where θ^* represents an external forcing function. For phase plane analysis, this may be either a step or ramp function, or zero. For definiteness, we will consider only the step input, in which case, the change of variable

$$x = \theta^* - \theta$$

converts Eq. (29) to

$$\ddot{x} + 2\zeta\omega\dot{x} + \omega^2 x = 0 \tag{30}$$

Putting this in the standard form

$$\begin{aligned} \dot{x} &= y \equiv P \\ \dot{y} &= -2\zeta\omega y - \omega^2 x \equiv Q \end{aligned} \tag{31}$$

we find that the singular point is at $x = y = 0$, and is characterized by

Stable Node:	$\zeta > 1, \ \omega > 0$	
Stable Focus:	$0 < \zeta < 1, \ \omega > 0$	
Center:	$\zeta = 0, \ \omega^2 > 0$	(32)
Unstable Node:	$\zeta < -1, \ \omega > 0$	
Unstable Focus:	$-1 < \zeta < 0, \ \omega > 0$	
Saddle:	$\omega^2 < 1$	

In the examples to follow, it will be more convenient to work with the θ - $\dot{\theta}$ plane rather than the x-y plane. The foregoing discussion indicates that all that is required is a translation of the origin. Consequently, one may work with Eq. (29) directly in the θ - $\dot{\theta}$ plane if one considers that in this case the singular point is at $\theta = \theta^*$, $\dot{\theta} = 0$.

We now consider the problem of a simplified pitch plane vehicle control system incorporating various types of nonlinearities.

Example 3: The system to be analyzed is shown in Fig. 15, with the control loop as depicted in Fig. 16. The symbols used have the following meaning:

I	=	moment of inertia of vehicle, slug-ft^2
K_A	=	servoamplifier gain, N. D.
K_R	=	rate gyro gain, sec
ℓ_c	=	control thrust moment arm, ft
ℓ_α	=	aerodynamic load moment arm, ft
L_α	=	aerodynamic load, lb/rad
s	$\equiv$	Laplace operator, sec^{-1}
T_c	=	control thrust, lb
t	=	time, sec
δ_c	=	engine command, deg
δ	=	engine angle, deg
θ	=	pitch angle, deg
θ_E	=	error signal, deg
θ_c	=	input signal, deg

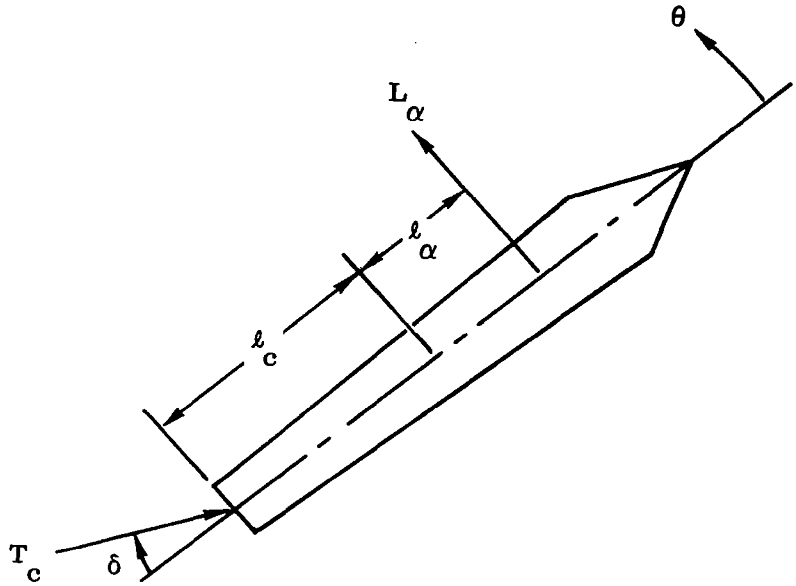


Figure 15. Vehicle Geometry for Example 3

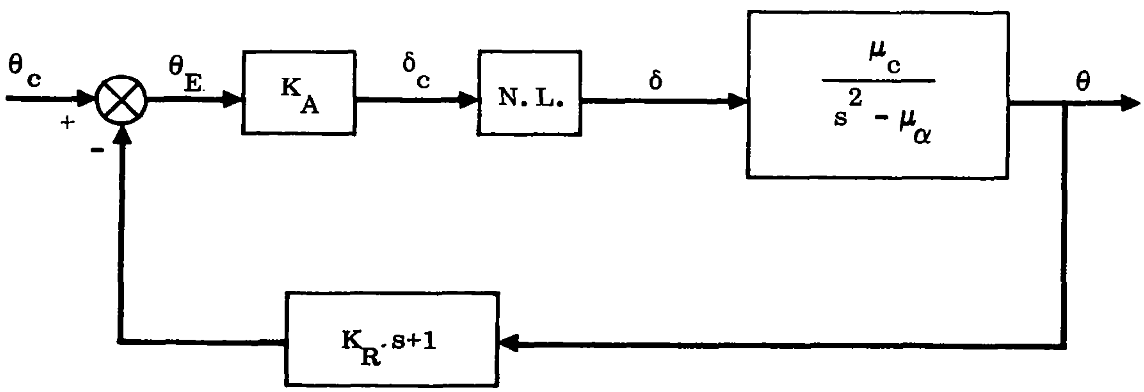


Figure 16. Control Loop for Example 3

$\mu_c = T_c \ell_c / I$ = control engine effectiveness parameter, $\sec^{-2}$

$\mu_\alpha = L_\alpha \ell_\alpha / I$ = aerodynamic effectiveness parameter, $\sec^{-2}$

The equations defining the motion are

$$\ddot{\theta} = \mu_c \delta + \mu_\alpha \theta \tag{33}$$

$$\delta_c = K_A \left[\theta_c - (K_R s + 1) \theta \right] \tag{34}$$

$$\begin{aligned} \theta_c &= A \quad t > 0 \\ &= 0 \quad t < 0 \end{aligned} \tag{35}$$

The nonlinearity is a saturation characterized by*

$$\begin{aligned} \delta &= B & \delta_c &> \delta_{CM} \\ &= \delta_c & -\delta_{CM} &\leqslant \delta_c \leqslant \delta_{CM} \\ &= -B & \delta_c &< -\delta_{CM} \end{aligned} \tag{36}$$

Dividing the $\dot{\delta}_c$ - δ_c phase plane into three regions as shown in Fig. 17, we may write the relevant equations as follows:

$$\text{Region } ① : \ddot{\delta}_c - \mu_\alpha \delta_c = \mu_\alpha \delta_1^* \tag{37}$$

$$\begin{aligned} \text{Region } ② : \ddot{\delta}_c + \mu_c K_A K_R \dot{\delta}_c + (\mu_c K_A - \mu_\alpha) \delta_c \\ = (\mu_c K_A - \mu_\alpha) \delta_2^* \end{aligned} \tag{38}$$

$$\text{Region } ③ : \ddot{\delta}_c - \mu_\alpha \delta_c = \mu_\alpha \delta_3^* \tag{39}$$

Here

$$\begin{aligned} \delta_1^* &= - \frac{K_A (\mu_\alpha A + \mu_c B)}{\mu_\alpha} \\ \delta_2^* &= - \frac{\mu_\alpha K_A A}{(\mu_c K_A - \mu_\alpha)} \\ \delta_3^* &= - \frac{K_A (\mu_\alpha A - \mu_c B)}{\mu_\alpha} \end{aligned} \tag{40}$$

* $\delta_{CM} \equiv$ saturation limit of δ_c

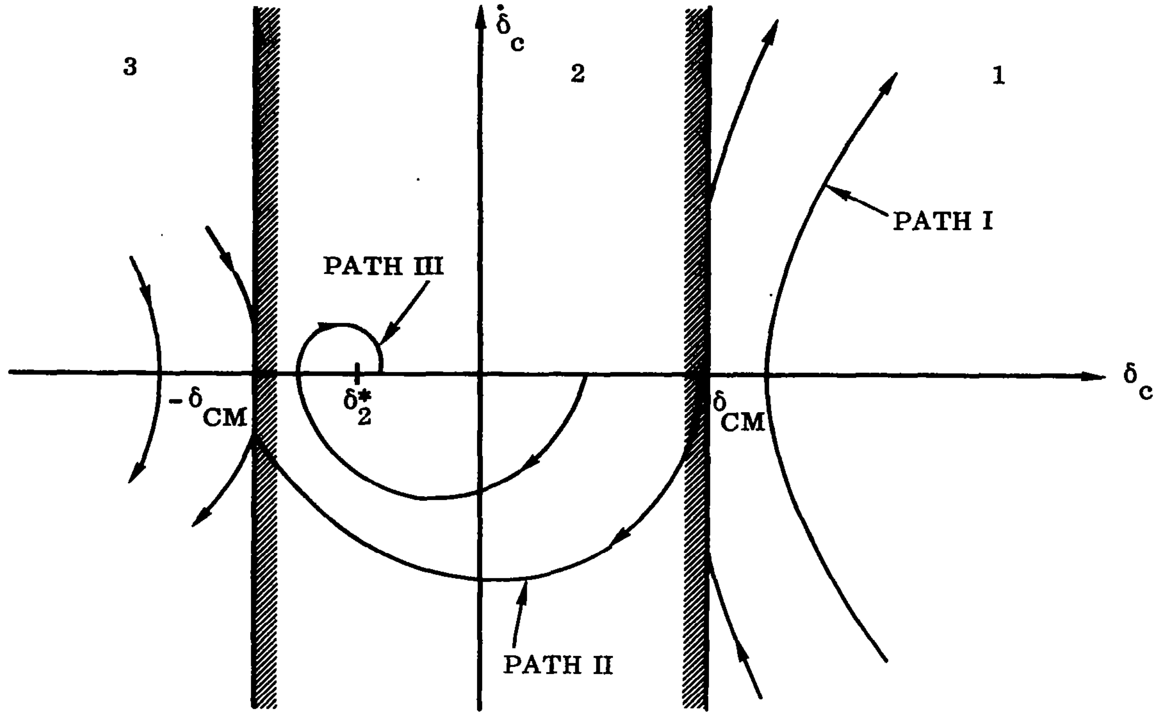


Figure 17. Phase Portrait for Example 3

The δ_i^*'s are the singular points for their respective regions. Note that a singular point need not be located inside its own region. When indeed it is located outside its own region it is known as a virtual singular point.

Let us assume that the system parameters are such that δ_1^* and δ_3^* are saddle points while δ_2^* is a stable focus. This situation exists when the parameters A, B, K_A, K_R, μ_c, μ_α, are all positive, $(\mu_c K_A - \mu_\alpha) > 0$, and the K_A and K_R are chosen appropriately. Then δ_2^* is located in region (2) as are also δ_1^* and δ_3^*. Typical trajectories for each region then have the form shown in Fig. 17. The initial conditions are $\delta_c = K_A A$, $\dot{\delta}_c = 0$; i.e., the trajectories all start on the positive δ_c axis. Now if $K_A A > \delta_{CM}$, then the system diverges; a typical form of the trajectory is as depicted by path I. If $K_A A < \delta_{CM}$, but with the magnitude of the step command such that path II is followed, then on entering region (3) the trajectory follows a divergent path, and the system is unstable. With the magnitude of the step input, A, reduced still further, a path such as III is followed and a point of stable equilibrium is reached with δ_c taking a steady state value equal to δ_2^*. There are thus only two possible types of response; one of which results in a stable equilibrium, and the other is an unbounded divergence. For given values of the system parameters, this is solely a function of the magnitude of the step input. For the case discussed herein (aerodynamically unstable vehicle), catastrophic instability is a potential problem.

In the case of an aerodynamically stable vehicle, the parameter μ_α is negative and the singular points, δ_1^* and δ_3^* are now centers. These will still be located within region (2) and the paths for regions (1) and (3) will be as shown in Fig. 18. Depending on the magnitude of the step input, the paths may or may not enter region (3) . In all events, an equilibrium condition will be reached with $\delta = \delta_2^*$ (which is now positive since μ_α is assumed negative).

Example 4: Taking again the case of the pitch plane autopilot, we investigate the influence of dead zone in the position gyro. The control loop is shown in Fig. 19a and the characteristics of the nonlinearity are depicted in Fig 19b. The relevant equations are:

$$\ddot{\theta} - \mu_\alpha \theta = \mu_c \delta$$

$$\delta = K_A (\theta_c - K_R \dot{\theta} - f)$$

$$f = \theta - \theta_o \qquad |\theta| > \theta_o$$

$$= 0 \qquad -\theta_o \leqslant \theta \leqslant \theta_o$$

θ_c is again taken to be a step input of magnitude, A. We may write therefore

$$\ddot{\theta} + \mu_c K_A K_R \dot{\theta} (\mu_c K_A - \mu_\alpha) \theta = (\mu_c K_A - \mu_\alpha) \theta_1^* \qquad \text{(a)}$$

when $|\theta| > \theta_o$, while

$$\ddot{\theta} + \mu_c K_A K_R \dot{\theta} - \mu_\alpha = - \mu_\alpha \theta_2^* \qquad \text{(b)}$$

if $|\theta| < \theta_o$, where

$$\theta_1^* = \frac{\mu_c K_A (A + \theta_o)}{(\mu_c K_A - \mu_\alpha)}$$

$$\theta_2^* = \frac{\mu_c K_A A}{\mu_\alpha}$$

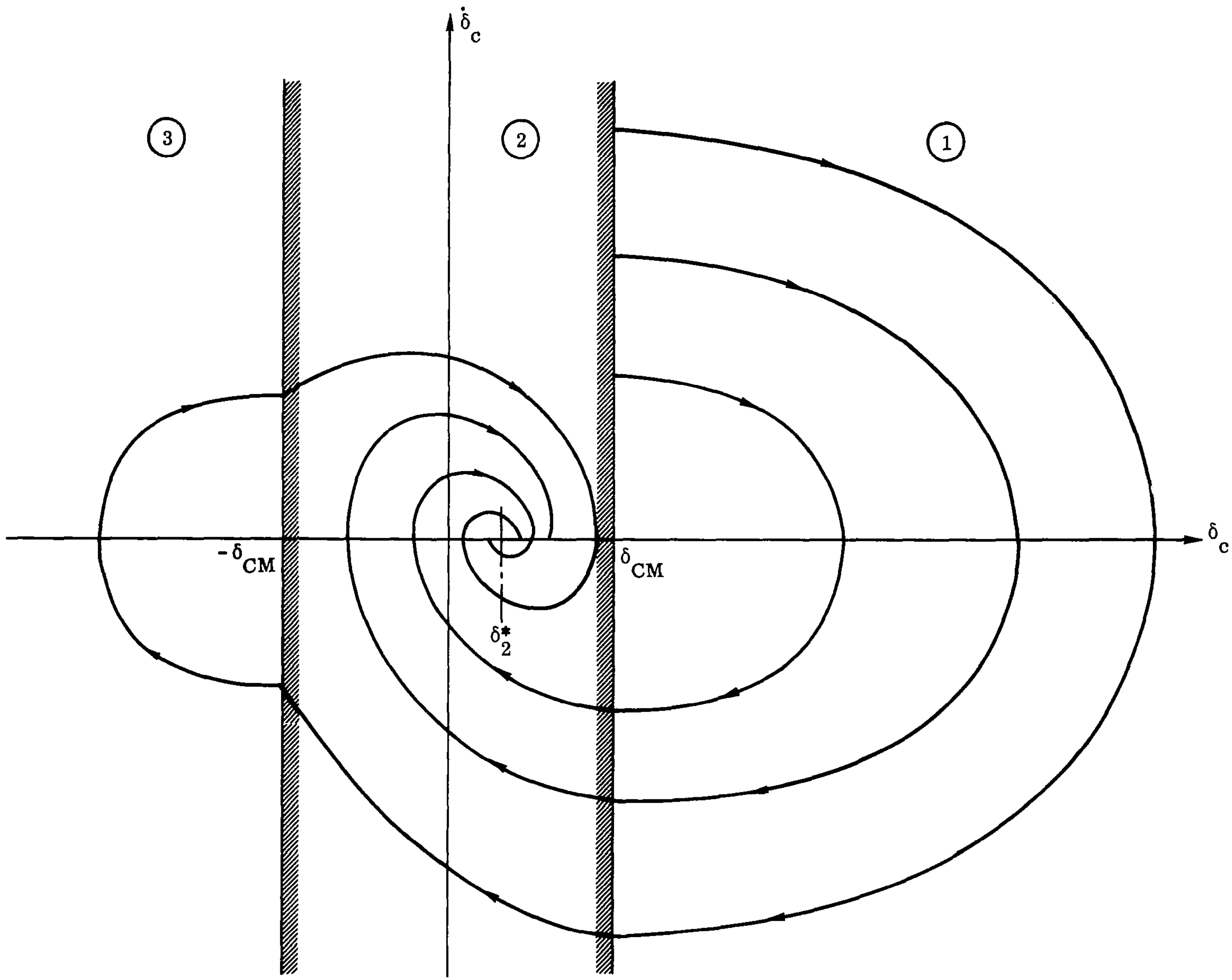
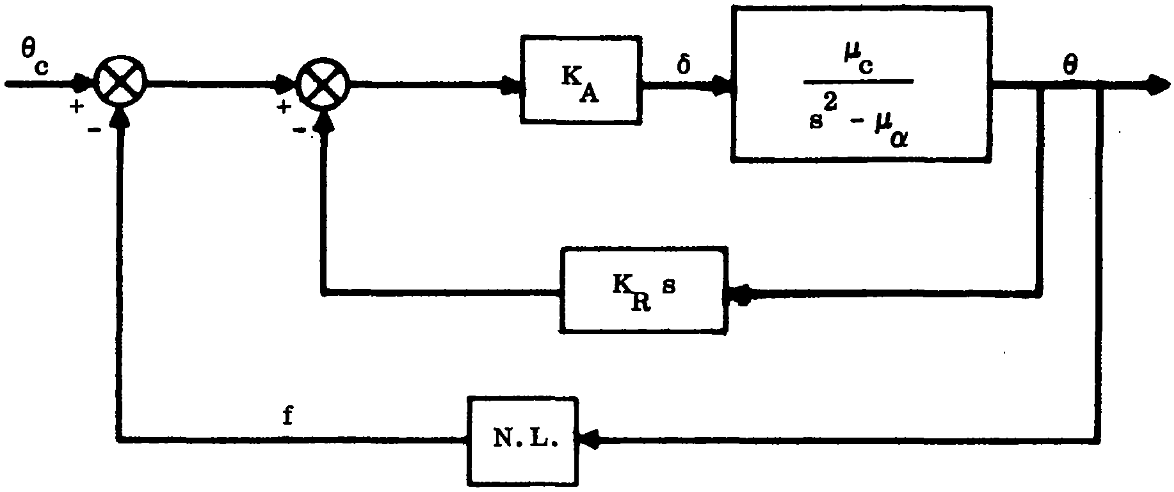
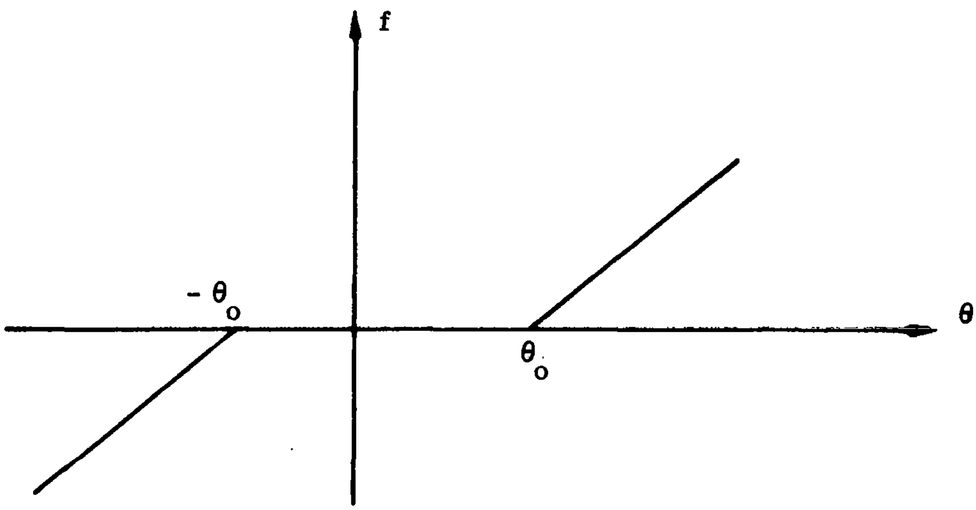


Figure 18. Phase Portrait for Example 3 (aerodynamically stable case)



a. CONTROL LOOP FOR EXAMPLE 4



b.

Figure 19. Characteristics of Nonlinear Element for Example 4

The phase plane is again divided into three regions as shown in Fig. 20. θ_1^* is the critical point for regions (1) and (3) while θ_2^* is the critical point for region (2) . When the quantity ($\mu_c K_A - \mu_\alpha$) is positive and the system gains are appropriately chosen, θ_1^* represents a stable focus, and θ_2^* is a saddle point. The trajectory starts at the origin ($\theta = \dot{\theta} = 0$ initially), and on reaching the region boundary, takes a path corresponding to the stable focus in region (1) . The critical point represents a point of stable equilibrium. Note that in the present case, with positive step input of magnitude A, θ is never negative.

Example 5: As a final illustration, consider the system shown in Fig. 21. The approach here is as follows. Damping must be provided to make the system stable for small errors; however, the response for large errors should be as fast as possible. This is accomplished by operating on the damping rather than the system gain. The damping, in fact, is made to depend on the system error rather than the output rate. The governing equations are as follows.

$$\delta = K_A \theta_E - f \dot{\theta}$$

$$\ddot{\theta} - \mu_\alpha \theta = \mu_c \delta$$

$$\theta_E = \theta_c - \theta$$

$$\begin{aligned} f &= K_A K_R \qquad & |\theta_E| \leqslant \theta_o \\ &= 0 \qquad & |\theta_E| > \theta_o \end{aligned}$$

In the $\theta_E - \dot{\theta}_E$ plane we have

$$\ddot{\theta}_E + \mu_c K_A K_R \dot{\theta}_E + (\mu_c K_A - \mu_\alpha)\theta_E = (\mu_c K_A - \mu_\alpha)\theta_E^*$$

for $|\theta_E| < \theta_o$

and

$$\ddot{\theta}_E + (\mu_c K_A - \mu_\alpha)\theta_E = (\mu_c K_A - \mu_\alpha)\theta_E^*$$

for $|\theta_E| > \theta_o$

where $\theta_E^* = - \dfrac{\mu_\alpha A}{\mu_c K_A - \mu_\alpha}$

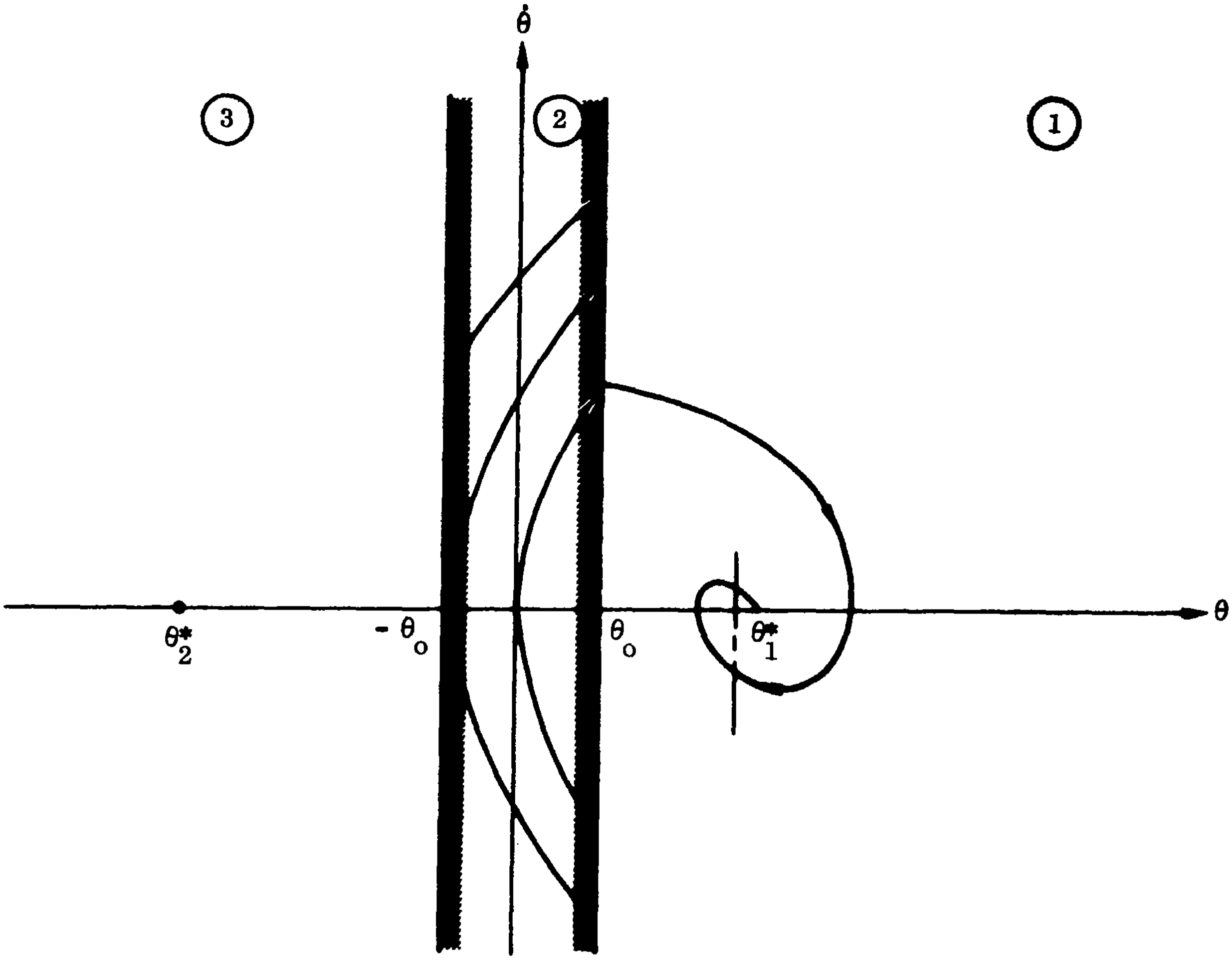


Figure 20. Phase Portrait for Example 4

We have again used a step input of magnitude, A. The location of the singular point is the same for all three regions as shown in Fig. 22. Assuming that the quantity ($\mu_c K_A - \mu_\alpha$) is positive, the singular point for regions 1 and 3 is a center, while for region 2 the singular point is either a stable focus or a stable node, depending on the system gain. For a sufficiently large step input, θ_E^* is located in region 3 as shown, and motion starts at the point $\theta_E = A$, $\dot{\theta}_E = 0$. In regions 1 and 3 the trajectories are ellipses with center θ_E^*, while in region 2 the trajectory is a spiral with the same center. There is a point of unstable equilibrium at $\dot{\theta}_E = 0$, $\theta_E = -\theta_o$. A slight disturbance will cause the motion to follow the trajectory of path I. It is obvious that for smaller values of step input such that θ_E^* is located in region 2 , the motion will reach a point of stable equilibrium.

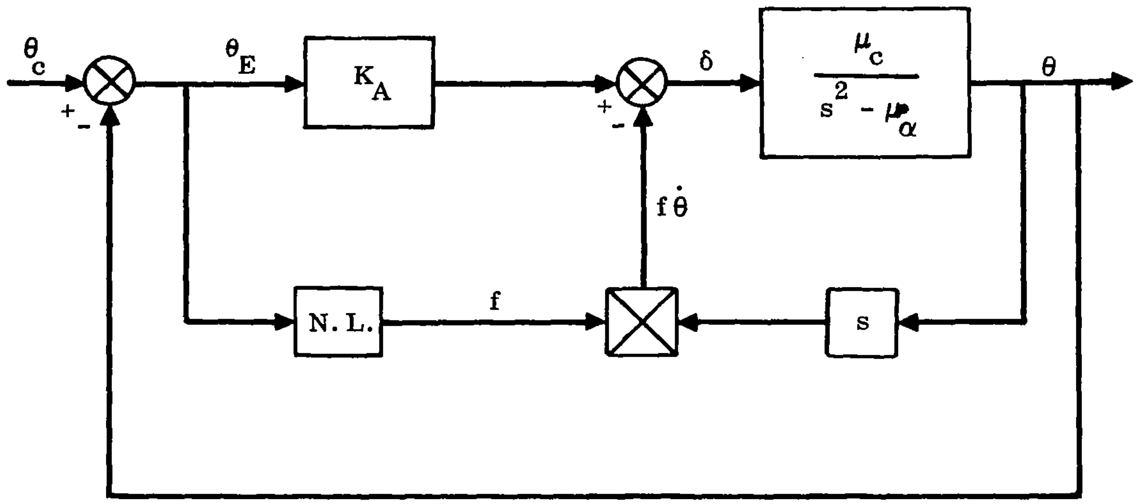


Figure 21. Control Loop of Example 5

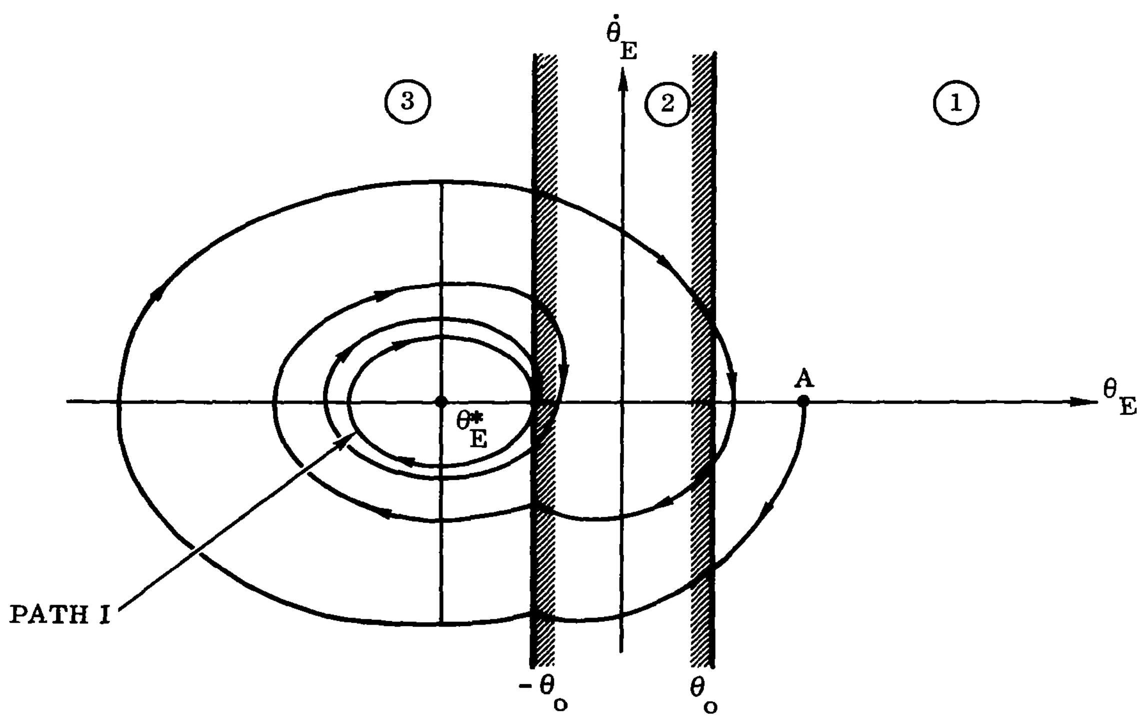


Figure 22. Phase Portrait for Example 5

3.2 DESCRIBING FUNCTIONS

Because of the fact that linear methods of analysis are so powerful and universally applicable, it is natural to seek to "linearize" the nonlinear systems encountered in control analysis. One well established technique involves the derivation of the (linear) equations characterizing small perturbations about steady-state conditions. When the motions are not small, this approach is invalid. The method of "describing functions," developed in this country by Kochenburger[8], may be used to analyze control systems incorporating one nonlinearity if the following conditions are satisfied:

a. The input to the nonlinear element is sinusoidal.

b. The output of the nonlinear element is periodic and of the same fundamental frequency as the input signal.

c. Only the fundamental of the output wave need be considered in frequency response analysis.

d. The nonlinear element is not time varying.

Under these conditions, the method yields an "equivalent linear gain" for the nonlinear element. Using frequency response methods, one is then able to determine the stability properties of the system, the frequency and amplitude of limit cycles* if they exist, and whether these limit cycles are stable or unstable.

In this section, no attempt will be made to develop the method in rigor or depth, since this topic is treated extensively in standard texts[9]. We will instead emphasize the basic principles and fundamental assumptions, and develop some very general forms for the describing function. Some recent developments such as the dual input describing function (DIDF) and methods for multiple nonlinearities will be discussed. Potential applications to launch vehicle control problems will be stressed.

3.2.1 Definition of the Describing Function

The conventional definition of a transfer function of a linear dynamic element is based on the response of the element to an input signal of the form

$$x = R_o \sin \omega t \tag{41}$$

The output is then

$$y = R_o' \sin (\omega t + \varphi) \tag{42}$$

*A limit cycle is an oscillation of fixed amplitude and frequency.

where in general R_0' is different from R_0 and φ is the angle of phase lag. The amplitude ratio, R_0'/R_0 and phase lag, φ, are functions of the input frequency but not of the input amplitude. If the dynamic element is not linear, then the response to a sinusoidal input is periodic but not sinusoidal. In seeking to define an equivalent transfer function for a nonlinear element, one is led to consider the frequency spectrum of the output and to assume that only the first harmonic is significant.

Let the input to the nonlinear element be given by Eq. (41). Then the output has the property

$$y(t + T) = y(t) \tag{43}$$

where $T = \frac{2\pi}{\omega}$. The output y (t) may be expanded in a Fourier series as follows:

$$y(t) = \frac{A_o}{2} + \sum_{n=1}^{\infty} A_n \cos \frac{2\pi nt}{T} + \sum_{n=1}^{\infty} B_n \sin \frac{2\pi nt}{T} \tag{44}$$

where

$$A_n = \frac{2}{T} \int_{\alpha}^{\alpha + T} y(t) \cos \left(\frac{2\pi nt}{T}\right) dt \qquad (45$$

$$B_n = \frac{2}{T} \int_{\alpha}^{\alpha + T} y(t) \sin \left(\frac{2\pi nt}{T}\right) dt \tag{46}$$

and α is any constant.

We note also

$$A_n \cos \frac{2\pi nt}{T} + B_n \sin \frac{2\pi nt}{T} = C_n \sin \frac{2\pi nt}{T} + \varphi_n$$

$$= C_n \sin \frac{2\pi nt}{T} \cos \varphi_n + C_n \cos \frac{2\pi nt}{T} \sin \varphi_n$$

Equating coefficients of like sine and cosine terms, we obtain

$$B_n = C_n \cos \varphi_n$$

$$A_n = C_n \sin \varphi_n$$

Solving this for C_n and φ_n gives

$$C_n = \sqrt{A_n^2 + B_n^2} \tag{47}$$

$$\varphi_n = \tan^{-1} \frac{A_n}{B_n} \tag{48}$$

Hence, Eq. 44 may be written in the form

$$y(t) = \frac{A_o}{2} + \sum_{n=1}^{\infty} C_n \sin\left(\frac{2\pi nt}{T} + \varphi_n\right) \tag{49}$$

Or

$$y(t) = \frac{A_o}{2} + \sum_{n=1}^{\infty} C_n \cos\left(\frac{2\pi nt}{T} - \frac{\pi}{2} + \varphi_n\right) \tag{50}$$

If we put

$$\alpha = -L$$

$$T = 2L$$

Then we have

$$A_n = \frac{1}{L} \int_{-L}^{L} y(t) \cos\left(\frac{\pi nt}{L}\right) dt$$

$$B_n = \frac{1}{L} \int_{-L}^{L} y(t) \sin\left(\frac{\pi nt}{L}\right) dt$$

We see, therefore, that if $y(t)$ is an even function [i.e., $y(t) = y(-t)$], then

$$B_n = 0$$

and

$$A_n = \frac{2}{L} \int_{o}^{L} y(t) \cos\left(\frac{\pi nt}{L}\right) dt$$

while if $y(t)$ is an odd function [i.e., $y(-t) = -y(t)$], then

$$A_n = 0$$

and

$$B_n = \frac{2}{L} \int_0^L y(t) \sin\left(\frac{\pi n t}{L}\right) dt$$

For the special case of

$$y(\omega t) = y(\omega t + 2\pi) \tag{51}$$

We have

$$y(\omega t) = \frac{A_o}{2} + \sum_{n=1}^{\infty} A_n \cos n\omega t + \sum_{n=1}^{\infty} B_n \sin n\omega t \tag{52}$$

where

$$A_n = \frac{1}{\pi} \int_{\alpha}^{\alpha + 2\pi} y(\omega t) \cos n\omega t \; d(\omega t) \tag{53}$$

$$B_n = \frac{1}{\pi} \int_{\alpha}^{\alpha + 2\pi} y(\omega t) \sin n\omega t \; d(\omega t) \tag{54}$$

If $y(\omega t) = y(-\omega t)$, then

$$B_n = 0$$

$$A_n = \frac{2}{\pi} \int_0^{\pi} y(\omega t) \cos n\omega t \; d(\omega t) \tag{55}$$

If $y(-\omega t) = -y(\omega t)$

$$A_n = 0$$

$$B_n = \frac{2}{\pi} \int_0^{\pi} y(\omega t) \sin n\omega t \; d(\omega t) \tag{56}$$

We now <u>define</u> the Describing Function as the complex ratio

$$G_D = \frac{C_1}{R_o} e^{j\varphi_1} \tag{57}$$

This may be expressed in the equivalent form

$$G_D = \frac{C_1}{R_o} \angle\varphi_1 = \frac{B_1}{R_o} + j\frac{A_1}{R_o}$$

$$= g_e(R_o) + j\, b_e(R_o) \tag{58}$$

As is obvious from the definition, the only significant quantity at the output of the nonlinear element is taken to be the fundamental harmonic of the Fourier spectrum. In this way, an equivalent linear transfer function (the describing function) is obtained. The validity of this approximation is strongly dependent on the low band-pass properties of the system considered. This condition is fulfilled in most systems of engineering interest and is, in fact, more accurate the higher the order of the system.

3.2.2 Describing Functions for Various Nonlinear Elements

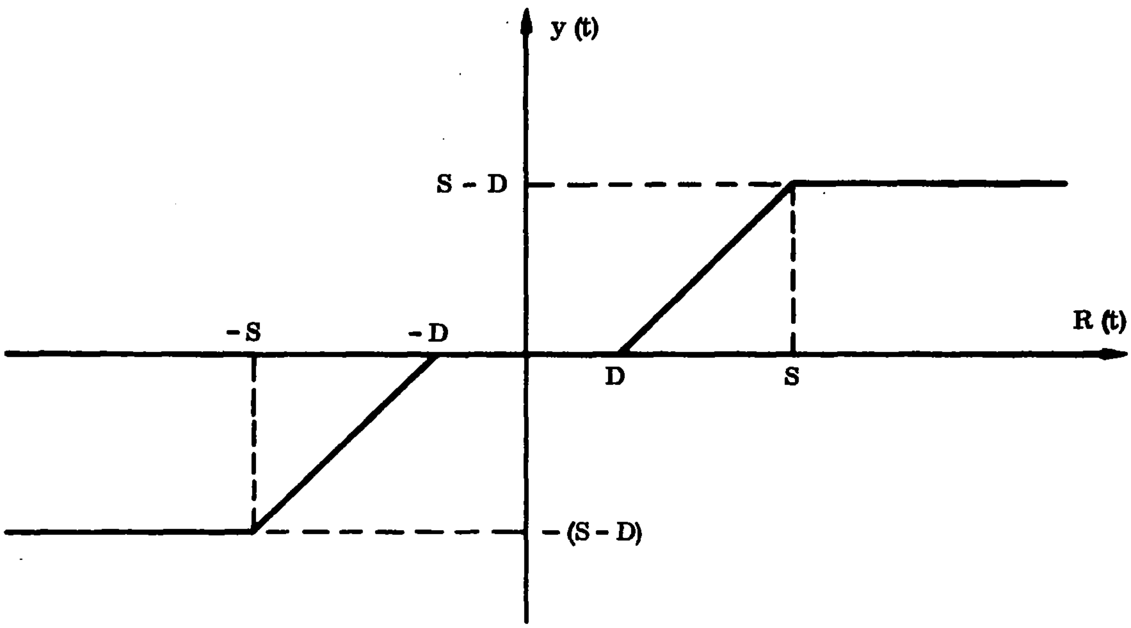
In this section, the describing function for various common types of nonlinearities will be developed. Following this, a very general describing function will be derived which is applicable to a wide class of piecewise linear elements. The section will conclude with an extensive tabulation of describing functions of the type most generally encountered in practice.

<u>Combined Saturation and Dead Zone</u>. This nonlinearity is illustrated in Fig. 23a, and the output waveform is shown in Fig. 23b. We have

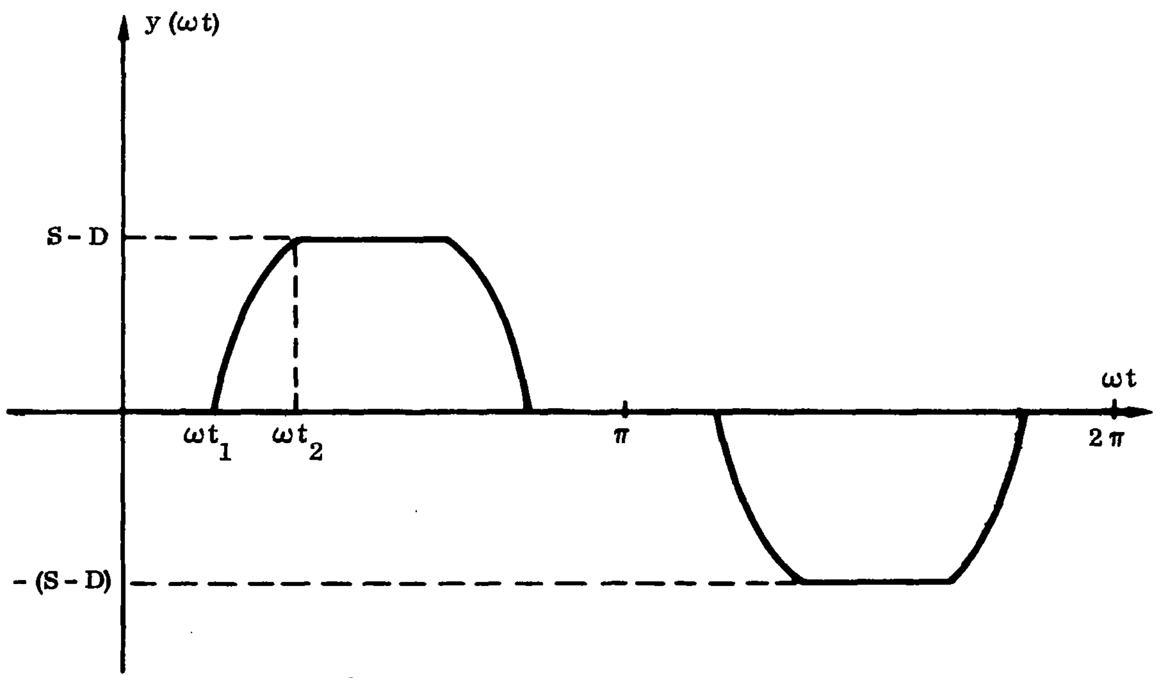
$$\begin{aligned} y(\omega t) &= 0 & 0 < \omega t < \omega t_1 \\ &= R_o(\sin\omega t - \sin\omega t_1) & \omega t_1 < \omega t < \omega t_2 \\ &= R_o(\sin\omega t_2 - \sin\omega t_1) & \omega t_2 < \omega t < \frac{\pi}{2} \end{aligned}$$

where

$$\sin\omega t_1 = {}^{D}/R_o \quad \text{and} \quad \sin\omega t_2 = {}^{S}/R_o$$



a. NONLINEAR ELEMENT



b. OUTPUT WAVEFORM

Figure 23. Combined Dead Zone and Saturation

Therefore

$$B_1 = \frac{4}{\pi} \int_{t_1}^{t_2} R_o \left(\sin \omega t - \sin \omega t_1 \right) \sin \omega t \; d(\omega t)$$

$$+ \frac{4}{\pi} \int_{t_2}^{\pi_2} R_o \left(\sin \omega t_2 - \sin \omega t_1 \right) \sin \omega t \; d(\omega t)$$

$$= \frac{2 R_o}{\pi} \left[\omega t_2 - \omega t_1 + \frac{\sin 2 \omega t_2}{2} - \frac{\sin 2 \omega t_1}{2} \right]$$

and

$$A_1 = 0$$

Hence

$$G_D = \frac{B_1}{R_o} = \frac{2}{\pi} \left[\omega t_2 - \omega t_1 + \frac{\sin 2 \omega t_2}{2} - \frac{\sin 2 \omega t_1}{2} \right] \tag{59}$$

The following special cases may be noted:

For Saturation only

$$\omega t_1 = 0$$

so that

$$G_D = \frac{2}{\pi} \left[\omega t_2 + \frac{\sin 2 \omega t_2}{2} \right] \tag{60}$$

And for Dead Zone only

$$\omega t_2 = \frac{\pi}{2}$$

so that

$$G_D = \frac{2}{\pi} \left[\frac{\pi}{2} - \omega t_1 - \frac{\sin 2 \omega t_1}{2} \right] \tag{61}$$

The describing function is shown graphically in Fig. 24.

Coulomb Friction. This case is shown in Fig. 25. We see that

$$y\ (\omega t) = -K \frac{R}{|R|} = -K \operatorname{sgn} R$$

or

$$\begin{aligned} y\ (\omega t) &= -K & R &> 0 \\ &= K & R &< 0 \end{aligned}$$

Hence

$$B_1 = \frac{4}{\pi} \int_0^{\pi_2} (-K) \sin \omega t \, d\,(\omega t) = -\frac{4K}{\pi}$$

$$A_1 = 0$$

so that

$$G_D = \frac{B_1}{R_o} = -\frac{4K}{\pi R_o} \tag{62}$$

Exponential Nonlinearity. In this case (Fig. 26), the nonlinearity is defined by:

$$y = |x^{n-1}|\, x$$

$$n > -2$$

Since y is an odd function,

$$A_n = 0$$

$$B_1 = \frac{2}{\pi} \int_0^{\pi} (R_o \sin \omega t)^n \sin \omega t \, d\,(\omega t)$$

$$= \frac{4 R_o^n}{\pi} \int_0^{\pi/2} (\sin \omega t)^{n+1} d\,(\omega t)$$

$$= \frac{2 R_o^n}{\sqrt{\pi}} \left[\frac{\Gamma \frac{n+2}{2}}{\Gamma \frac{n+3}{2}} \right]$$

where Γ ( ) ≡ Gamma Function. Hence

$$G_D = \frac{2 R_o^{n-1}}{\sqrt{\pi}} \left[\frac{\Gamma \frac{n+2}{2}}{\Gamma \frac{n+3}{2}} \right] \tag{63}$$

Relay with Dead Zone and Hysteresis. The essential elements of this nonlinearity are depicted in Fig. 27. Here we find that

$$A_o = \frac{K}{\pi} \int_{\omega \alpha}^{\pi - \omega\beta} d(\omega t) - \frac{K}{\pi} \int_{\pi + \omega\alpha}^{2\pi - \omega\beta} d(\omega t) = 0$$

$$A_1 = \frac{K}{\pi} \int_{\omega\alpha}^{\pi - \omega\beta} \cos \omega t \, d(\omega t) - \frac{K}{\pi} \int_{\pi + \omega\alpha}^{2\pi - \omega\beta} \cos \omega t \, d(\omega t)$$

$$= -\frac{2K}{\pi} (\sin \omega\alpha - \sin \omega\beta)$$

$$B_1 = \frac{K}{\pi} \int_{\omega\alpha}^{\pi - \omega\beta} \sin \omega t \, d(\omega t) - \frac{K}{\pi} \int_{\pi + \omega\alpha}^{2\pi - \omega\beta} \sin \omega t \, d(\omega t)$$

$$= \frac{2K}{\pi} (\cos \omega\alpha + \cos \omega\beta)$$

where

$$\sin \omega\alpha = \frac{\Delta + h}{2 R_o}$$

$$\sin \omega\beta = \frac{\Delta - h}{2 R_o}$$

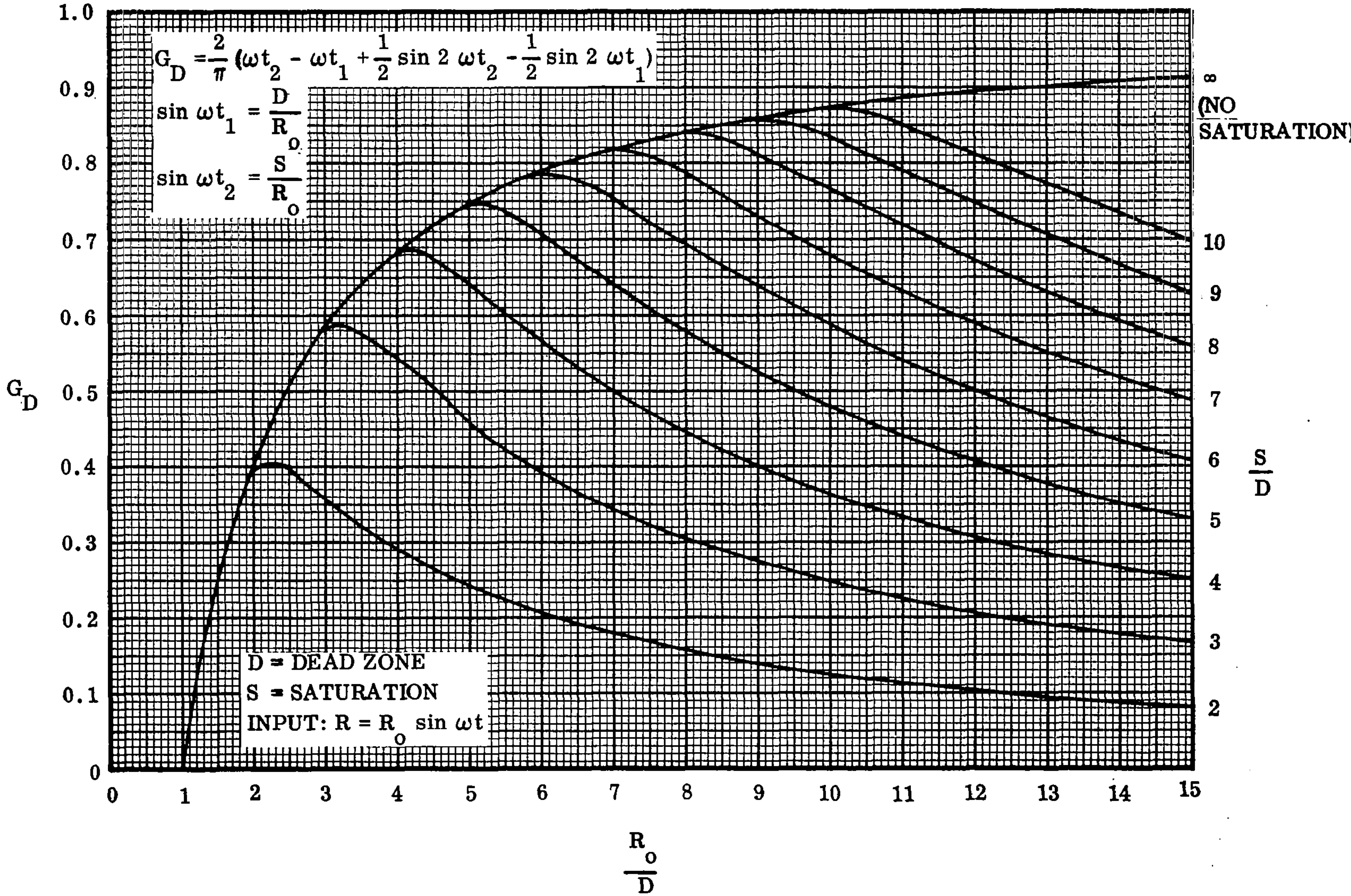
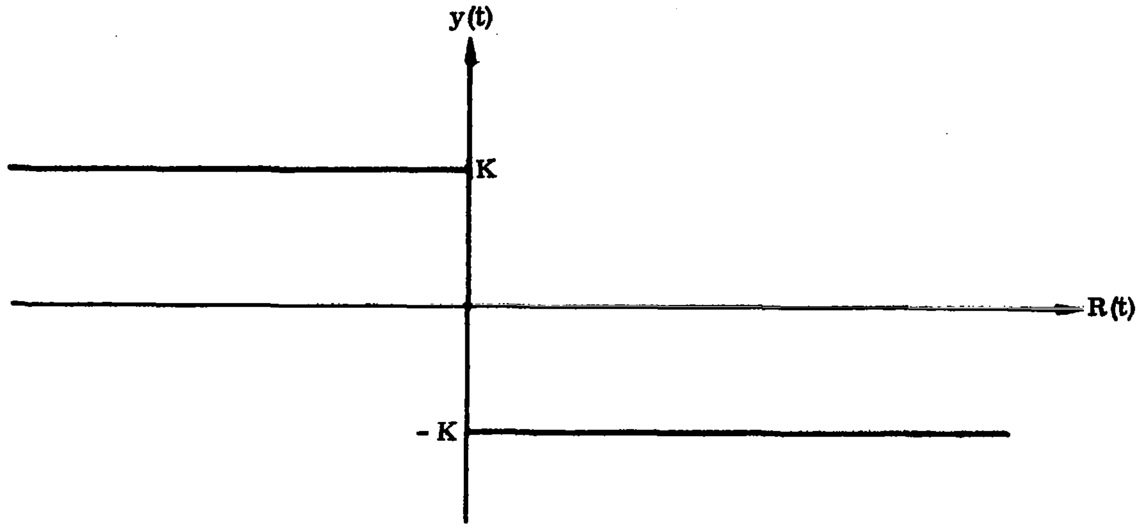
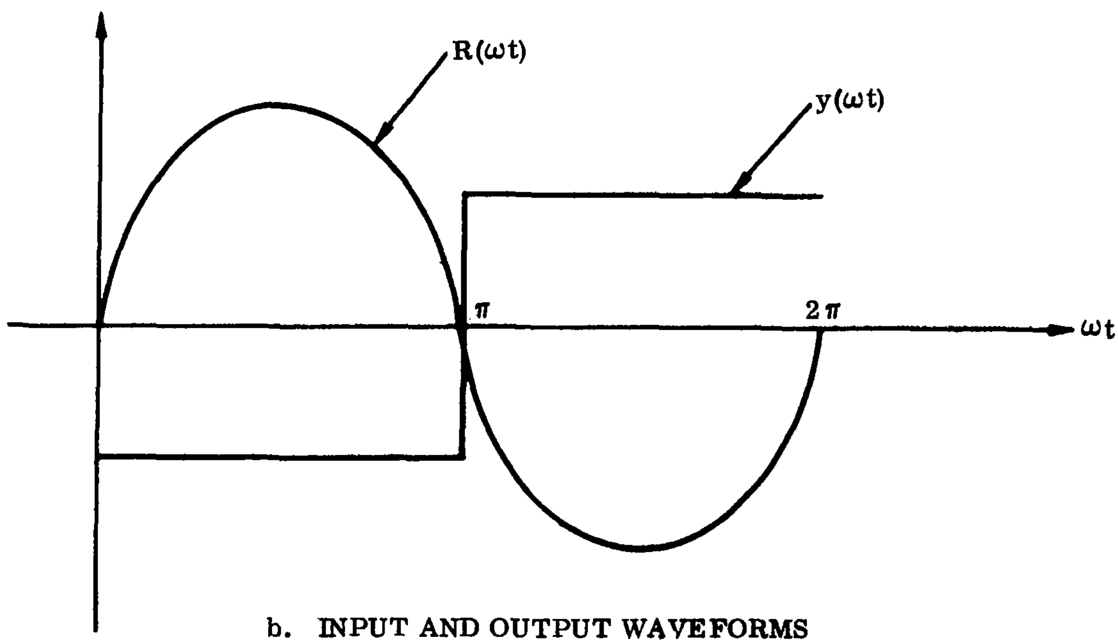


Figure 24. Describing Function for Saturation and Dead Zone

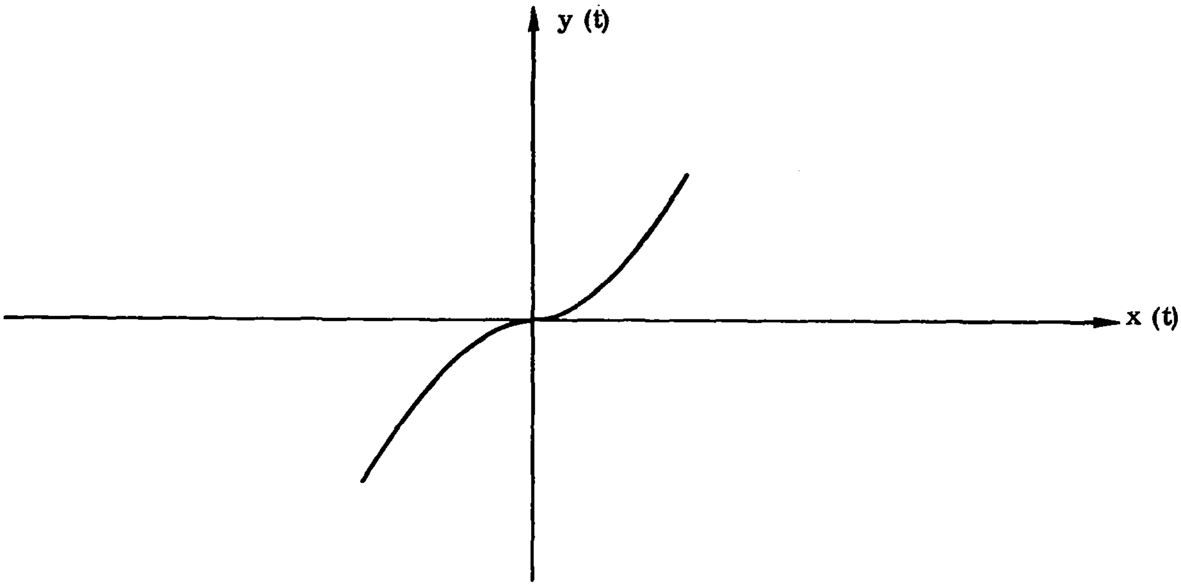


a. NONLINEAR ELEMENT

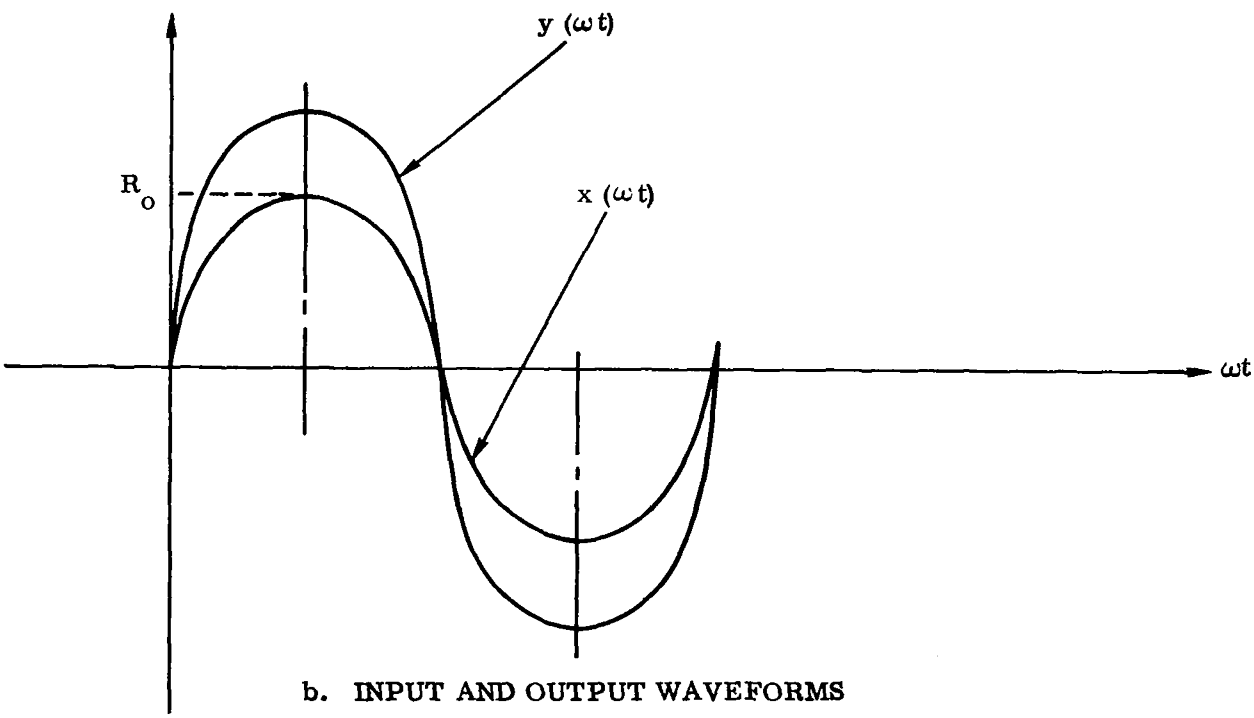


b. INPUT AND OUTPUT WAVEFORMS

Figure 25. Coulomb Friction

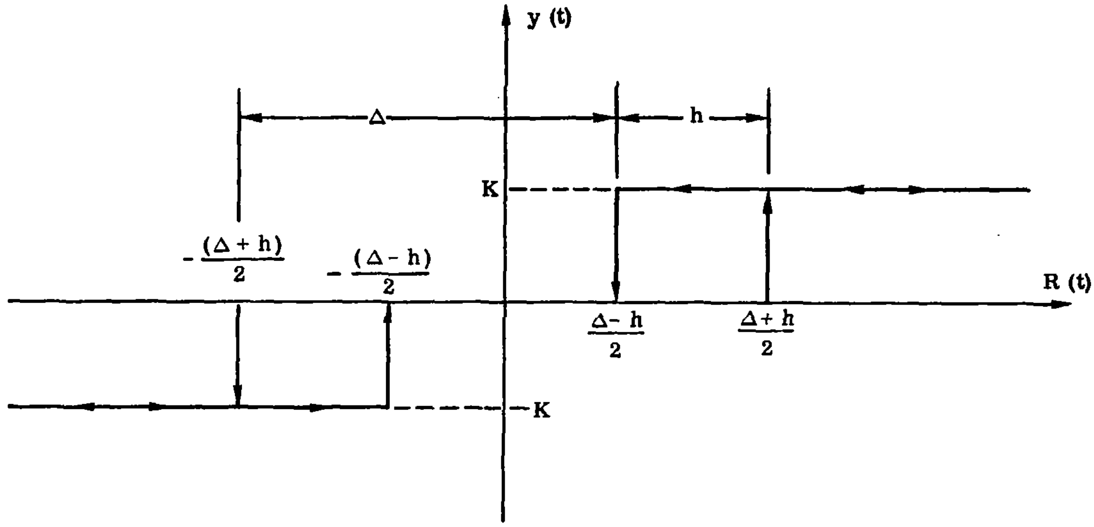


a. NONLINEAR ELEMENT

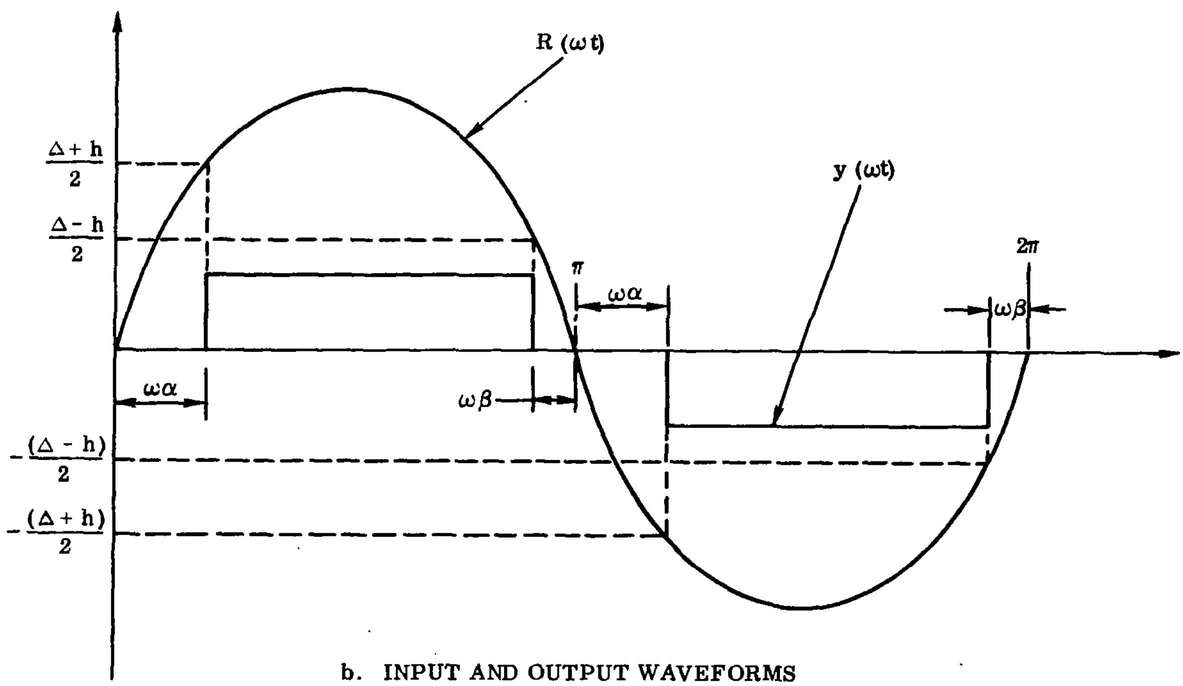


b. INPUT AND OUTPUT WAVEFORMS

Figure 26. Exponential Nonlinearity



a. NONLINEAR ELEMENT



b. INPUT AND OUTPUT WAVEFORMS

Figure 27. Relay with Dead Zone and Hysteresis

Now

$$C_1 = \sqrt{A_1^2 + B_1^2}$$

$$= \frac{4K}{\pi} \cos \left[\frac{\omega(\alpha + \beta)}{2}\right]$$

and

$$\varphi_1 = \tan^{-1} \frac{A_1}{B_1}$$

$$= -\frac{\omega}{2} (\alpha - \beta)$$

Putting

$$\psi = \omega\alpha$$

$$\theta = \omega\beta$$

we find

$$G_D = \frac{C_1}{R_o} \angle\varphi_1$$

$$= \frac{4K}{\pi R_o} \cos\left(\frac{\psi + \theta}{2}\right) \angle -\frac{1}{2}(\psi - \theta) \tag{64}$$

Figs. 28 through 30 depict this describing function graphically. In these figures, a normalized ordinate is used; viz., $G_D' = \frac{\Delta}{K} G_D$.

This reduces to

a. <u>No hysteresis</u> (h = 0)

$$G_D = \frac{4K}{\pi R_o} \sqrt{1 - \left(\frac{\Delta}{2 R_o}\right)^2} \tag{65}$$

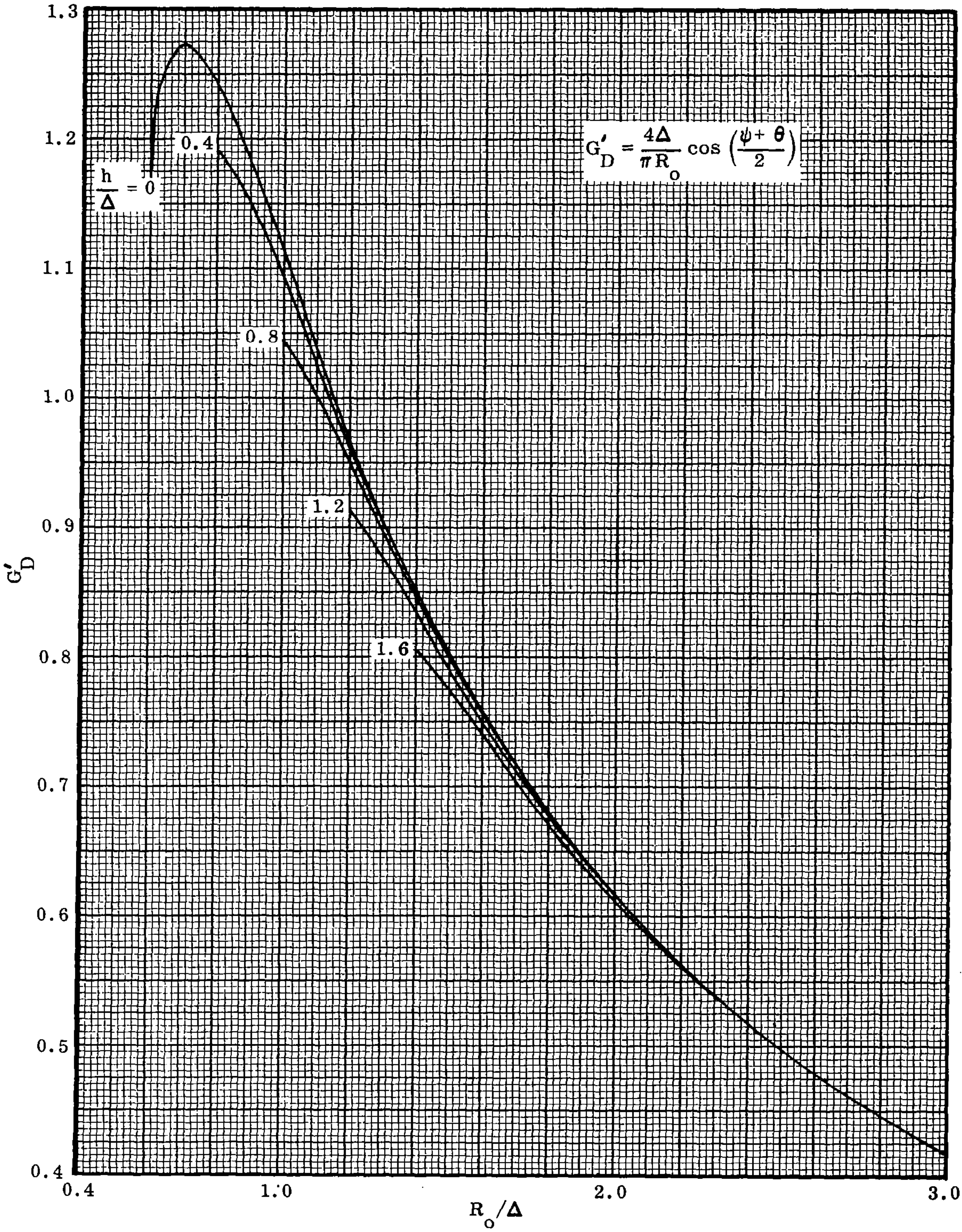


Figure 28. Relay Describing Function

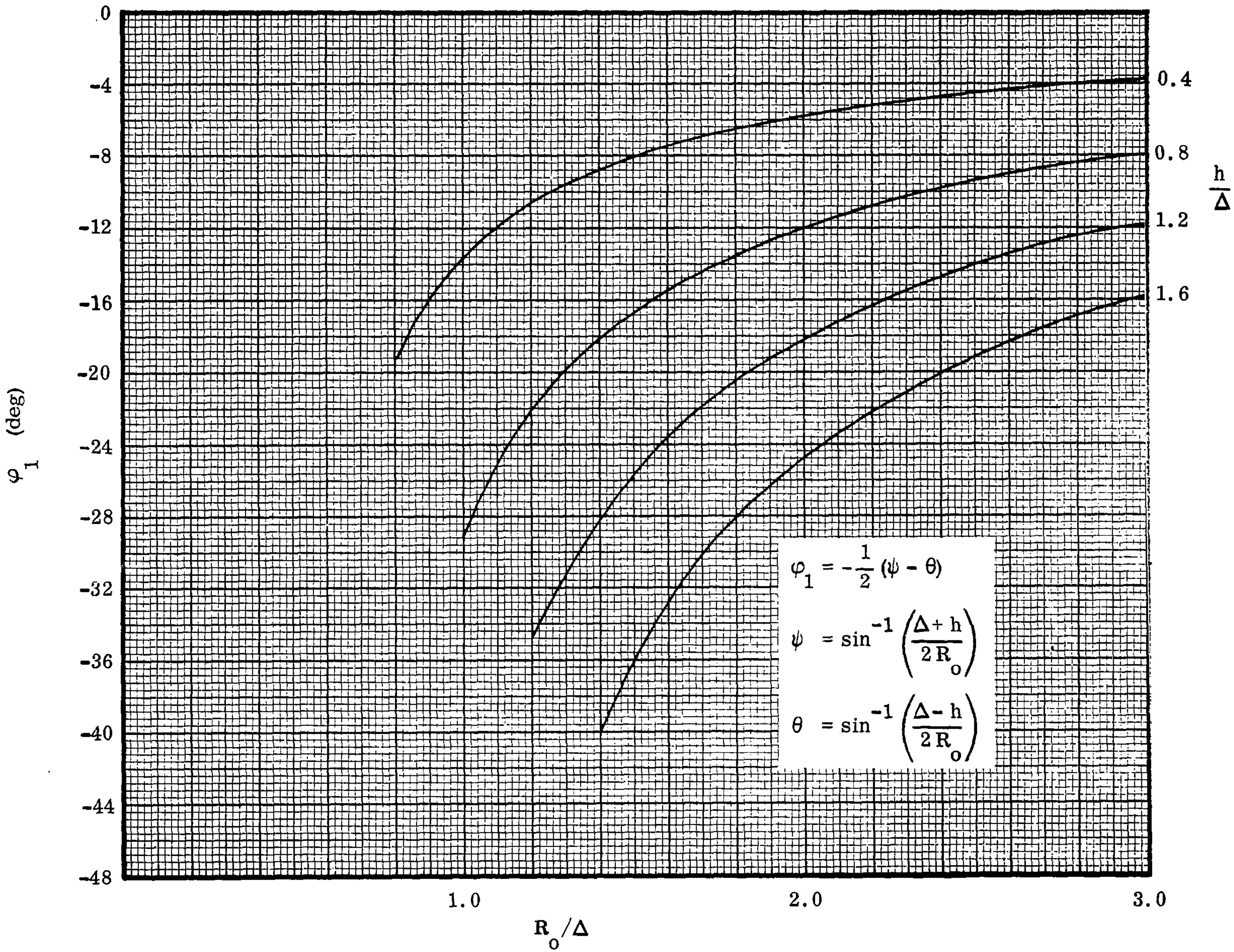


Figure 29. Relay Describing Function Phase Angle

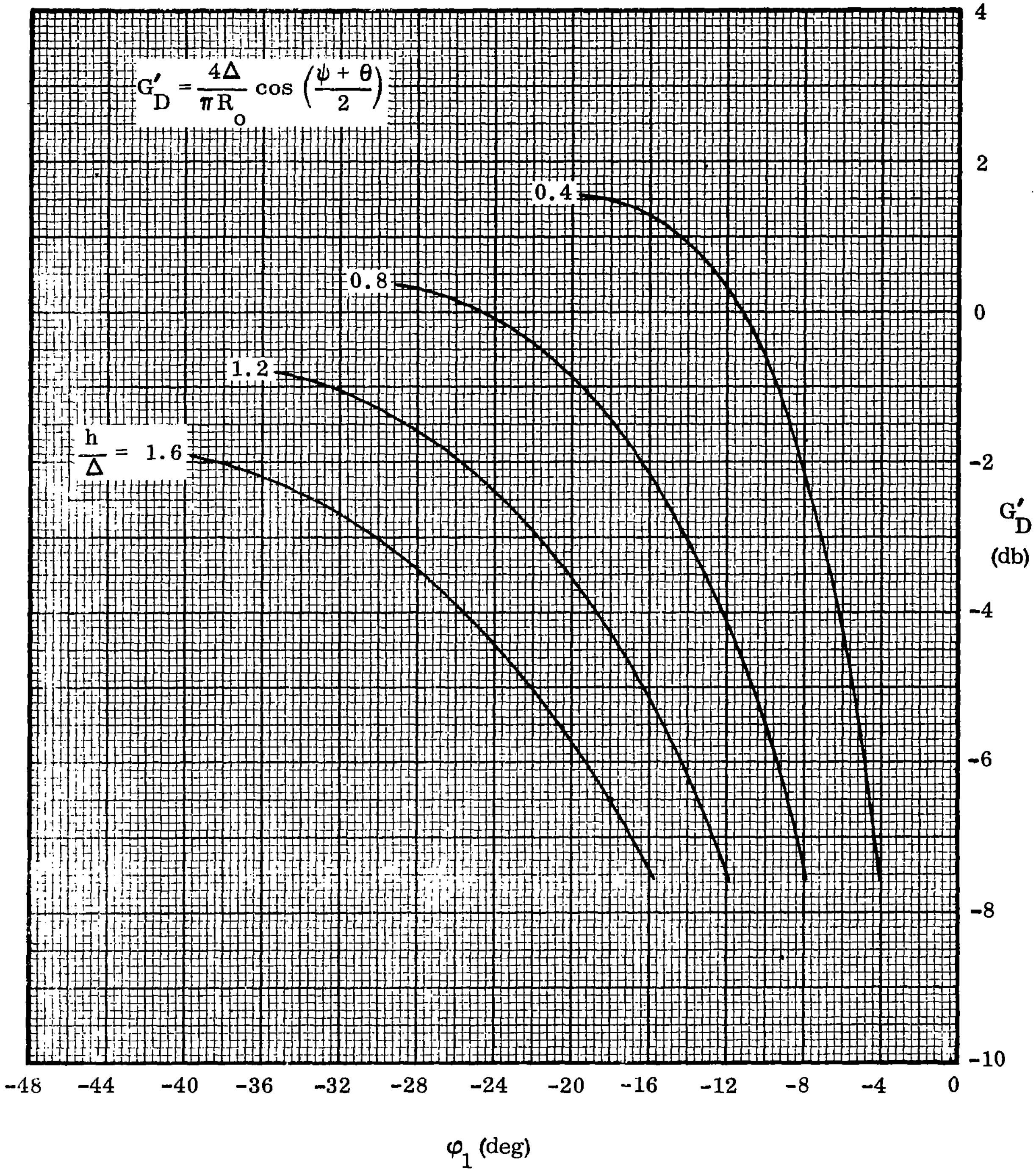


Figure 30. Relay Describing Function, Amplitude versus Phase

b. No dead zone $(\Delta = 0)$

$$G_D = \frac{4K}{\pi R_o} \angle -\delta \tag{66}$$

$$\delta = \sin^{-1}\left(\frac{h}{2 R_o}\right)$$

c. No hysteresis or dead zone $(h = \Delta = 0)$

$$G_D = \frac{4K}{\pi R_o} \tag{67}$$

Hysteresis. The following relations are evident from Fig. 31

$$
\begin{aligned}
y(\omega t) &= R_o \sin \omega t - \frac{a}{2} && 0 < \omega t < \frac{\pi}{2} \\
&= R_o - \frac{a}{2} && \frac{\pi}{2} < \omega t < \pi - \omega t_1 \\
&= R_o \sin \omega t + \frac{a}{2} && \pi - \omega t_1 < \omega t < \frac{3\pi}{2} \\
&= -\left(R_o - \frac{a}{2}\right) && \frac{3\pi}{2} < \omega t < 2\pi - \omega t_1 \\
&= R_o \sin \omega t - \frac{a}{2} && 2\pi - \omega t_1 < \omega t < 2\pi
\end{aligned}
$$

where

$$\sin \omega t_1 = \frac{R_o - a}{R_o}$$

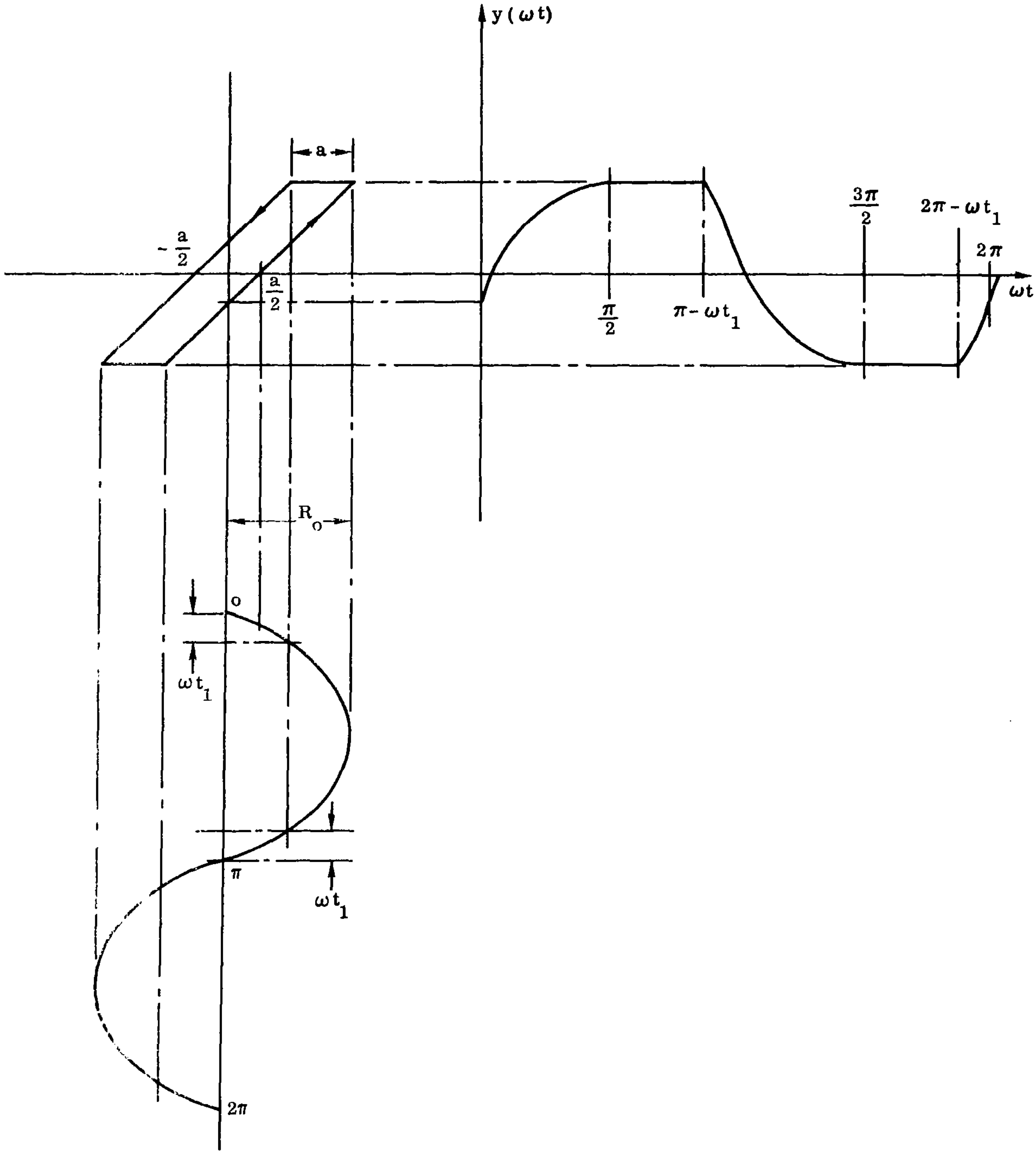


Figure 31. Hysteresis

Therefore

$$A_1 = \frac{R_o}{\pi} \int_{o}^{\frac{\pi}{2}} \sin \omega t \cos \omega t \, d(\omega t) - \frac{a}{2\pi} \int_{o}^{\frac{\pi}{2}} \cos \omega t \, d(\omega t)$$

$$+ \frac{R_o}{\pi} \int_{\frac{\pi}{2}}^{\pi - \omega t_1} \cos \omega t \, d(\omega t) - \frac{a}{2\pi} \int_{\frac{\pi}{2}}^{\pi - \omega t_1} \cos \omega t \, d(\omega t)$$

$$+ \frac{R_o}{\pi} \int_{\pi - \omega t_1}^{\frac{3\pi}{2}} \sin \omega t \cos \omega t \, d(\omega t) + \frac{a}{2\pi} \int_{\pi - \omega t_1}^{\frac{3\pi}{2}} \cos \omega t \, d(\omega t)$$

$$- \frac{R_o}{\pi} \int_{\frac{3\pi}{2}}^{2\pi - \omega t_1} \cos \omega t \, d(\omega t) + \frac{a}{2\pi} \int_{\frac{3\pi}{2}}^{2\pi - \omega t_1} \cos \omega t \, d(\omega t)$$

$$+ \frac{R_o}{\pi} \int_{2\pi - \omega t_1}^{2\pi} \sin \omega t \cos \omega t \, d(\omega t) - \frac{a}{2\pi} \int_{2\pi - \omega t_1}^{2\pi} \cos \omega t \, d(\omega t)$$

$$= \frac{a}{\pi} \left[\frac{a}{R_o} - 2 \right]$$

$$
\begin{aligned}
B_1 = {} & \frac{R_o}{\pi} \int_{0}^{\frac{\pi}{2}} \sin^2 \omega t \, d\,(\omega t) - \frac{a}{2\pi} \int_{0}^{\frac{\pi}{2}} \sin \omega t \, d\,(\omega t) \\
& + \frac{R_o}{\pi} \int_{\frac{\pi}{2}}^{\pi - \omega t_1} \sin \omega t \, d\,(\omega t) - \frac{a}{2\pi} \int_{\frac{\pi}{2}}^{\pi - \omega t_1} \sin \omega t \, d\,(\omega t) \\
& + \frac{R_o}{\pi} \int_{\pi - \omega t_1}^{\frac{3\pi}{2}} \sin^2 \omega t \, d\,(\omega t) + \frac{a}{2\pi} \int_{\pi - \omega t_1}^{\frac{3\pi}{2}} \sin \omega t \, d\,(\omega t) \\
& - \frac{R_o}{\pi} \int_{\frac{3\pi}{2}}^{2\pi - \omega t_1} \sin \omega t \, d\,(\omega t) + \frac{a}{2\pi} \int_{\frac{3\pi}{2}}^{2\pi - \omega t_1} \sin \omega t \, d\,(\omega t) \\
& + \frac{R_o}{\pi} \int_{2\pi - \omega t_1}^{2\pi} \sin^2 \omega t \, d\,(\omega t) - \frac{a}{2\pi} \int_{2\pi - \omega t_1}^{2\pi} \sin \omega t \, d\,(\omega t) \\
= {} & \frac{R_o}{\pi} \left[\frac{\pi}{2} + \omega t_1 + 1 - \left(\frac{a}{R_o} \right) \cos \omega t_1 \right]
\end{aligned}
$$

$$A_o = \frac{1}{\pi}\int_0^{\pi}\left[R_o \sin\omega t - \frac{a}{2}\right] d(\omega t) + \frac{1}{\pi}\int_{\frac{\pi}{2}}^{\pi-\omega t_1}\left[R_o - \frac{a}{2}\right] d(\omega t)$$

$$+ \frac{1}{\pi}\int_{\pi-\omega t_1}^{\frac{3\pi}{2}}\left[R_o \sin\omega t + \frac{a}{2}\right] d(\omega t) - \frac{1}{\pi}\int_{\frac{3\pi}{2}}^{2\pi-\omega t_1}\left[R_o - \frac{a}{2}\right] d(\omega t)$$

$$+ \frac{1}{\pi}\int_{2\pi-\omega t_1}^{2\pi}\left[R_o \sin\omega t - \frac{a}{2}\right] d(\omega t) = 0$$

and we have

$$G_D = \frac{\sqrt{A_1^2 + B_1^2}}{R_o} \angle \tan^{-1}\frac{A_1}{B_1}$$

$$= \frac{1}{\pi}\left[\frac{a^2}{R_o^2}\left(\frac{a}{R_o} - 2\right)^2 + \left(\frac{\pi}{2} + \omega t_1 + \sin\omega t_1 \cos\omega t_1\right)^2\right]^{\frac{1}{2}} \angle \tan^{-1}\beta \tag{68}$$

where

$$\beta = \frac{a\left(\frac{a}{R_o} - 2\right)}{R_o\left(\frac{\pi}{2} + \omega t_1 + \sin\omega t_1 \cos\omega t_1\right)}$$

A General Describing Function for Piecewise Linear Elements. Most of the describing functions derived thus far are special cases of a more general form[10] shown in Fig. 32. Applying Eqs. (53), (54), and (58) to this case, we find

$$g(R_o) = \frac{n_1}{\pi R_o}\left[R_o(-\theta_1 + 2\theta_2 + \theta_3 - \theta_4 - \theta_5)\right.$$

$$+ \frac{R_o}{2}(\sin 2\theta_1 - 2\sin 2\theta_2 - \sin 2\theta_3 + \sin 2\theta_4 + \sin 2\theta_5)$$

$$\left. + 2a(-\cos\theta_1 + 2\cos\theta_2 - \cos\theta_5) + 2d(\cos\theta_3 - \cos\theta_4)\right] \tag{69}$$

$$+ \frac{n_2}{\pi R_o}\left[R_o(\theta_5 - \theta_3) + \frac{R_o}{2}(\sin 2\theta_3 - \sin 2\theta_5)\right.$$

$$\left. + 2k_2(\cos\theta_5 - \cos\theta_3)\right]$$

$$+ \frac{n_3}{\pi R_o}\left[R_o(\pi - 2\theta_2) + R_o \sin 2\theta_2 - 4k_3\cos\theta_2\right]$$

and

$$b(R_o) = \frac{n_1}{\pi R_o}\left[\frac{R_o}{2}(\cos 2\theta_1 + \cos 2\theta_3 - \cos 2\theta_4\right.$$

$$\left. - \cos 2\theta_5) + 2a(\sin\theta_1 - \sin\theta_5) + 2d(\sin\theta_3 - \sin\theta_4)\right] \tag{70}$$

$$+ \frac{n_2}{\pi R_o}\left[\frac{R_o}{2}(-\cos 2\theta_3 + \cos 2\theta_5) + 2k_2(\sin\theta_5 - \sin\theta_3)\right]$$

where

$$k_2 = \frac{Dc - Fe}{D - F}$$

$$k_3 = b - \frac{B}{n_3}$$

$$\theta_1 = \sin^{-1}\frac{a}{R_o} \qquad \theta_2 = \sin^{-1}\frac{b}{R_o} \qquad \theta_3 = \sin^{-1}\frac{c}{R_o}$$

$$\theta_4 = \sin^{-1}\frac{d}{R_o} \qquad \theta_5 = \sin^{-1}\frac{e}{R_o}$$

Notice that the θ_i's have been defined such that they all lie in the first quadrant. A variety of special cases may be deduced from Eqs. (69) and (70). An extensive tabulation of these is given in Table 1, which is adapted from Ref. 10.

A General Polynomial Type Nonlinearity. The input-output characteristic of this type of element may be expressed by the equation

$$y = c_n x^n + c_{n-1} x^{n-2} |x| + c_{n-2} x^{n-2} + c_{n-3} x^{n-4} |x| + \ldots\ldots\ldots$$

$$\ldots + c_2 x |x| + c_1 x$$

There is no loss of generality in assuming that n is odd. The absolute value signs are used to ensure that y is an odd function of x. Proceeding in a manner completely analogous to that used in deriving Eq. (63), we find

$$G_D = \frac{2}{\sqrt{\pi}} \sum_{k=1}^{n} c_k R_o^{k-1} \frac{\Gamma\left(\frac{k+2}{2}\right)}{\Gamma\left(\frac{k+3}{2}\right)} \tag{71}$$

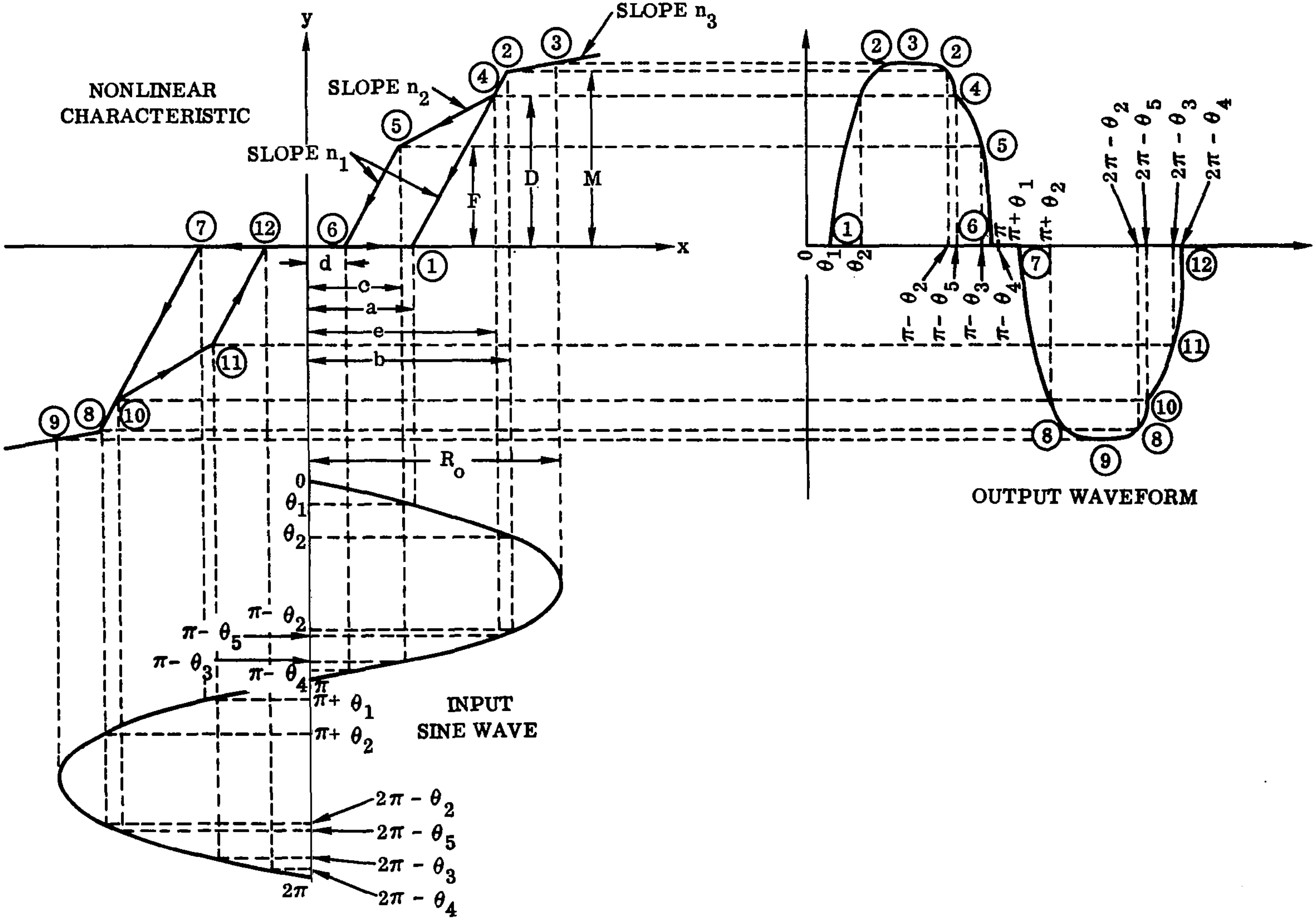


Figure 32. A General Piecewise Linear Nonlinearity

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements

Nonlinear devices with memory

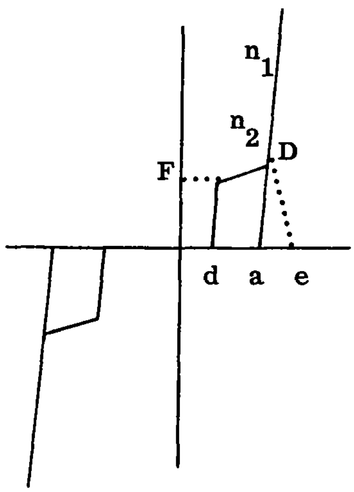
Identifying classification	Characteristic	Describing function
1		$g(R_o) = \frac{n_1}{\pi R_o}\Big[R_o(\pi - \theta_1 + \theta_3 - \theta_4 - \theta_5)$ $+ \frac{R_o}{2}(\sin 2\theta_1 - \sin 2\theta_3 + \sin 2\theta_4 + \sin 2\theta_5)$ $- 2a(\cos\theta_1 + \cos\theta_5) + 2d(\cos\theta_3 - \cos\theta_4)\Big]$ $+ \frac{n_2}{\pi R_o}\Big[R_o(\theta_5 + \theta_3) + \frac{R_o}{2}(\sin 2\theta_3 - \sin 2\theta_5)$ $+ 2k_2(\cos\theta_5 - \cos\theta_3)\Big]$ $b(R_o) = \frac{n_1}{\pi R_o}\Big[\frac{R_o}{2}(\cos 2\theta_1 + \cos 2\theta_3 - \cos 2\theta_4$ $- \cos 2\theta_5)$ $+ 2a(\sin\theta_1 - \sin\theta_5) + 2d(\sin\theta_3 - \sin\theta_4)\Big]$ $+ \frac{n_2}{\pi R_o}\Big[\frac{R_o}{2}(\cos 2\theta_5 - \cos 2\theta_3)$ $+ 2k_2(\sin\theta_5 - \sin\theta_3)\Big]$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

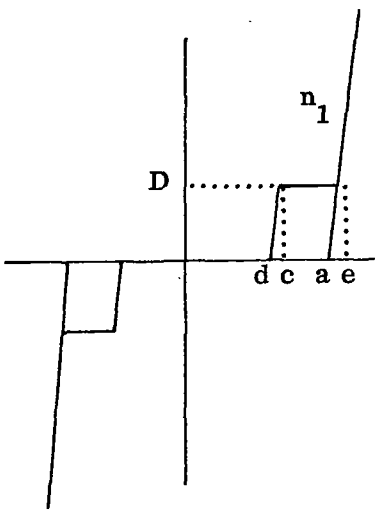
Identifying classification	Characteristic	Describing function
2	 	$g(R_o) = \frac{n_1}{\pi R_o}\Big[R_o(\pi - \theta_1 + \theta_3 - \theta_4 - \theta_5) + \frac{R_o}{2}(\sin 2\theta_1 - \sin 2\theta_3 + \sin 2\theta_4 + \sin 2\theta_5) - 2a(\cos\theta_1 + \cos\theta_5) + 2d(\cos\theta_3 - \cos\theta_4)\Big] - \frac{2D}{\pi R_o}(\cos\theta_5 - \cos\theta_3)$ $b(R_o) = \frac{n_1}{\pi R_o}\Big[\frac{R_o}{2}(\cos 2\theta_1 + \cos 2\theta_3 - \cos 2\theta_4 - \cos 2\theta_5) + 2a(\sin\theta_1 - \sin\theta_5) + 2d(\sin\theta_3 - \sin\theta_4)\Big] + \frac{2D}{\pi R_o}(\sin\theta_3 - \sin\theta_5)$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

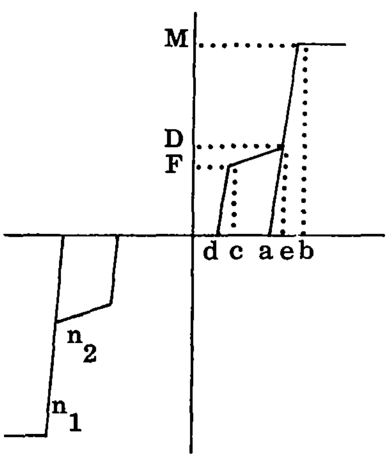
Identifying classification	Characteristic	Describing function
3		$g(R_o) = \frac{n_1}{\pi R_o}\Big[R_o(-\theta_1 + 2\theta_2 + \theta_3 - \theta_4 - \theta_5)$ $+ \frac{R_o}{2}(\sin 2\theta_1 - 2\sin 2\theta_2 - \sin 2\theta_3$ $+ \sin 2\theta_4 + \sin 2\theta_5)$ $+ 2a(-\cos\theta_1 + 2\cos\theta_2 - \cos\theta_5)$ $+ 2d(\cos\theta_3 - \cos\theta_4)\Big]$ $+ \frac{n_2}{\pi R_o}\Big[R_o(\theta_5 - \theta_3) + \frac{R_o}{2}(\sin 2\theta_3 - \sin 2\theta_5)$ $+ 2k_2(\cos\theta_5 - \cos\theta_3)\Big] + \frac{4M}{\pi R_o}\cos\theta_2$ $b(R_o) = \frac{n_1}{\pi R_o}\Big[\frac{R_o}{2}(\cos 2\theta_1 + \cos 2\theta_3 - \cos 2\theta_4$ $- \cos 2\theta_5)$ $+ 2a(\sin\theta_1 - \sin\theta_5) + 2d(\sin\theta_3 - \sin\theta_4)\Big]$ $+ \frac{n_2}{\pi R_o}\Big[\frac{R_o}{2}(-\cos 2\theta_3 + \cos 2\theta_5)$ $+ 2k_2(\sin\theta_5 - \sin\theta_3)\Big]$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

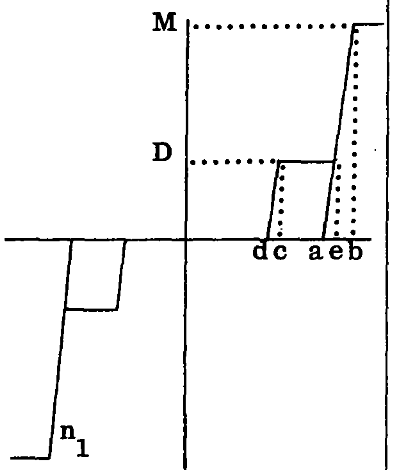
Identifying classification	Characteristic	Describing function
4		$g(R_o) = \frac{n_1}{\pi R_o} \Big[R_o(-\theta_1 + 2\theta_2 + \theta_3 - \theta_4 - \theta_5)$ $+ \frac{R_o}{2} (\sin 2\theta_1 - 2 \sin 2\theta_2$ $- \sin 2\theta_3 + \sin 2\theta_4 + \sin 2\theta_5)$ $+ 2a(-\cos\theta_1 + 2\cos\theta_2 - \cos\theta_5) + 2d(\cos\theta_3$ $- \cos\theta_4) \Big] + \frac{2D}{\pi R_o}(\cos\theta_3 - \cos\theta_5) + \frac{4M}{\pi R_o}\cos\theta_2$ $b(R_o) = \frac{n_1}{\pi R_o} \Big[\frac{R_o}{2} (\cos 2\theta_1 + \cos 2\theta_3 - \cos 2\theta_4$ $- \cos 2\theta_5) + 2a(\sin\theta_1 - \sin\theta_5) + 2d(\sin\theta_3$ $- \sin\theta_4) \Big] + \frac{2D}{\pi R_o}(\sin\theta_3 - \sin\theta_5)$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

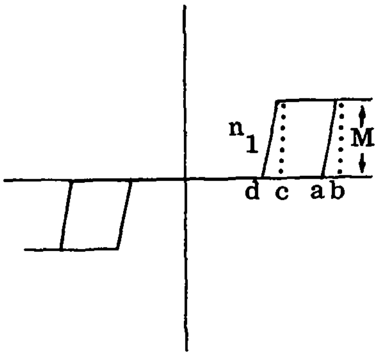
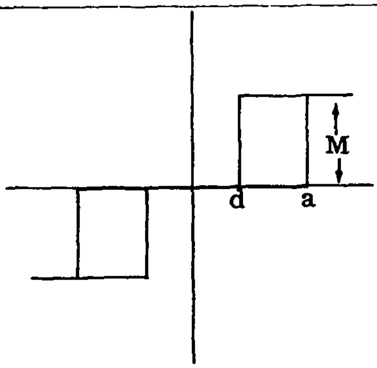
Identifying classification	Characteristic	Describing function
5		$g(R_o) = \frac{n_1}{\pi R_o}\Big[R_o(-\theta_1 + \theta_2 + \theta_3 - \theta_4) + \frac{R_o}{2}(\sin 2\theta_1 - \sin 2\theta_2 - \sin 2\theta_3 + \sin 2\theta_4) + 2a(-\cos\theta_1 + \cos\theta_2) + 2d(\cos\theta_3 - \cos\theta_4)\Big] + \frac{2M}{\pi R_o}(\cos\theta_2 + \cos\theta_3)$ $b(R_o) = \frac{n_1}{\pi R_o}\Big[\frac{R_o}{2}(\cos 2\theta_1 - \cos 2\theta_2 + \cos 2\theta_3 - \cos 2\theta_4) + 2a(\sin\theta_1 - \sin\theta_2) + 2d(\sin\theta_3 - \sin\theta_4)\Big] - \frac{2M}{\pi R_o}(\sin\theta_2 - \sin\theta_3)$
6		$g(R_o) = \frac{2M}{\pi R_o}(\cos\theta_1 + \cos\theta_3)$ $b(R_o) = \frac{-2M}{\pi R_o}(\sin\theta_1 - \sin\theta_3)$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

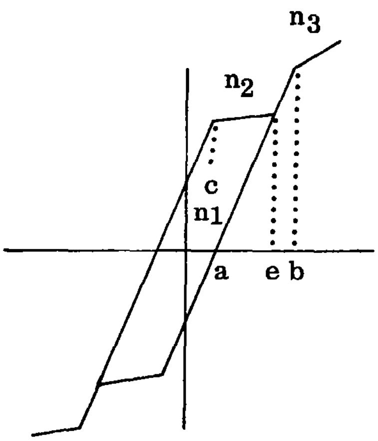
Identifying classification	Characteristic	Describing function
7		$g(R_o) = \frac{n_1}{\pi R_o}\Big[R_o(2\theta_2 + \theta_3 - \theta_5) + \frac{R_o}{2}(-2\sin 2\theta_2 - \sin 2\theta_3 + \sin 2\theta_5) + 2a(2\cos\theta_2 - \cos\theta_3 - \cos\theta_5)\Big] + \frac{n_2}{\pi R_o}\Big[R_o(\theta_5 - \theta_3) + \frac{R_o}{2}(\sin 2\theta_3 - \sin 2\theta_5) + 2k_2(\cos\theta_5 - \cos\theta_3)\Big] + \frac{n_3}{\pi R_o}\Big[R_o(\pi - 2\theta_2) + R_o\sin 2\theta_2 - 4k_3\cos\theta_2\Big]$ $b(R_o) = \frac{n_1}{\pi R_o}\Big[\frac{R_o}{2}(\cos\theta_3 - \cos 2\theta_5) + 2a(\sin\theta_3 + \sin\theta_5)\Big] + \frac{n_2}{\pi R_o}\Big[\frac{R_o}{2}(-\cos 2\theta_3 + \cos 2\theta_5) + 2k_2(\sin\theta_5 - \sin\theta_3)\Big]$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

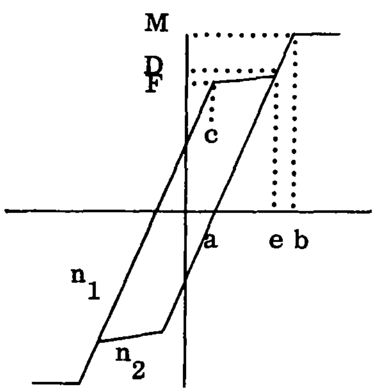
Identifying classification	Characteristic	Describing function
8		$\begin{aligned} g(R_o) &= \frac{n_1}{\pi R_o}\Big[R_o(2\theta_2 + \theta_3 - \theta_5) + \frac{R_o}{2} -2\sin 2\theta_2 \\ &- \sin 2\theta_3 + \sin 2\theta_5) + 2a(2\cos\theta_2 - \cos\theta_3 \\ &- \cos\theta_5)\Big] + \frac{n_2}{\pi R_o}\Big[R_o(\theta_5 - \theta_3) + \frac{R_o}{2}\sin 2\theta_3 \\ &- \sin 2\theta_5) + 2k_2(-\cos\theta_3 + \cos\theta_5)\Big] \\ &+ \frac{4M}{\pi R_o}\cos\theta_2 \end{aligned}$ $\begin{aligned} b(R_o) &= \frac{n_1}{\pi R_o}\Big[\frac{R_o}{2}(\cos 2\theta_3 - \cos 2\theta_5) \\ &-2a(\sin\theta_3 + \sin\theta_5)\Big] + \frac{n_2}{\pi R_o}\Big[\frac{R_o}{2}(-\cos 2\theta_3 \\ &+ \cos 2\theta_5) + 2k_2(\sin\theta_5 - \sin\theta_3)\Big] \end{aligned}$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

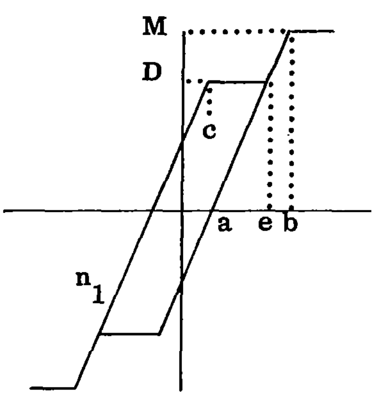
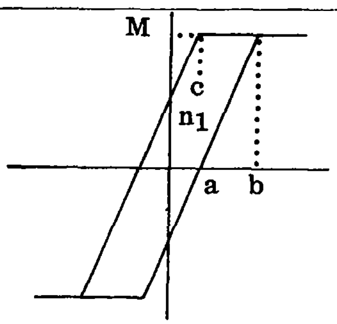
Identifying classification	Characteristic	Describing function
9		$g(R_o) = \frac{n_1}{\pi R_o}\left[R_o(2\theta_2 + \theta_3 - \theta_5) + \frac{R_o}{2}(-2\sin 2\theta_2 - \sin 2\theta_3 + \sin 2\theta_5) + 2a(2\cos\theta_2 - \cos\theta_3 - \cos\theta_5)\right] + \frac{2D}{\pi R_o}(\cos\theta_3 - \cos\theta_5) + \frac{4M}{\pi R_o}\cos\theta_2$ $b(R_o) = \frac{n_1}{\pi R_o}\left[\frac{R_o}{2}(\cos 2\theta_3 - \cos 2\theta_5) - 2a(\sin\theta_3 + \sin\theta_5)\right] + \frac{2D}{\pi R_o}(\sin\theta_3 - \sin\theta_5)$
10		$g(R_o) = \frac{n_1}{\pi R_o}\left[R_o(\theta_2 + \theta_3) + \frac{R_o}{2}(-\sin 2\theta_2 - \sin 2\theta_3) + 2a(\cos\theta_2 - \cos\theta_3)\right] + \frac{2M}{\pi R_o}(\cos\theta_2 + \cos\theta_3)$. $b(R_o) = \frac{n_1}{\pi R_o}\left[\frac{R_o}{2}(\cos 2\theta_3 - \cos 2\theta_2) - 2a(\sin\theta_2 + \sin\theta_3)\right] + \frac{2M}{\pi R_o}(\sin\theta_3 - \sin\theta_2)$ $R_o > b$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

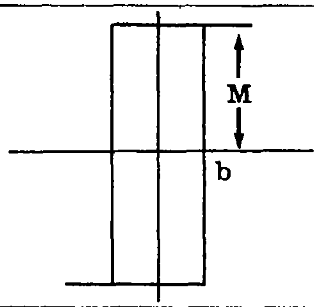
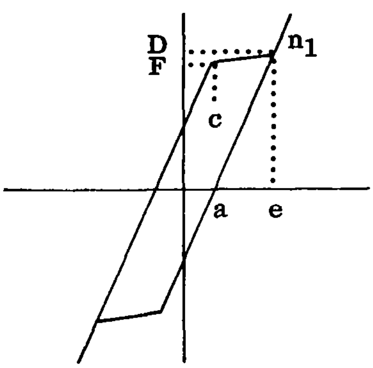
Identifying classification	Characteristic	Describing function
11		$g(R_o) = \frac{4M}{\pi R_o} \cos \theta_2$ $b(R_o) = \frac{-4M}{\pi R_o} \sin \theta_2$
12		$g(R_o) = \frac{n_1}{\pi R_o} \left[R_o (\pi + \theta_2 - \theta_5) + \frac{R_o}{2} (-\sin 2\theta_3 + \sin 2\theta_5) - 2a (\cos \theta_3 + \cos \theta_5) \right] + \frac{n_2}{\pi R_o} \left[R_o (\theta_5 - \theta_3) + \frac{R_o}{2} (\sin 2\theta_3 - \sin 2\theta_5) + 2k_2 (\cos \theta_5 - \cos \theta_3) \right]$ $b(R_o) = \frac{n_1}{\pi R_o} \left[\frac{R_o}{2} (\cos 2\theta_3 - \cos 2\theta_5) - 2a (\sin \theta_3 + \sin \theta_5) \right] + \frac{n_2}{\pi R_o} \left[\frac{R_o}{2} (-\cos 2\theta_3 + \cos 2\theta_5) + 2k_2 (\sin \theta_5 - \sin \theta_3) \right]$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

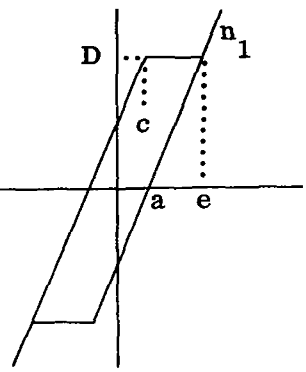
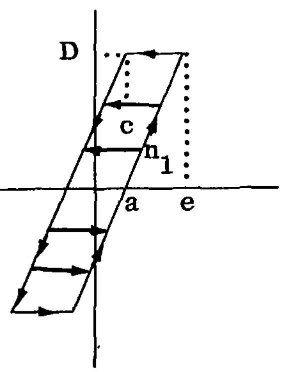
Identifying classification	Characteristic	Describing function
13		$g(R_o) = \frac{n_1}{\pi R_o}\left[R_o(\pi + \theta_3 - \theta_5) + \frac{R_o}{2}(-\sin 2\theta_3 + \sin 2\theta_5) - 2a(\cos\theta_3 + \cos\theta_5)\right] + \frac{2D}{\pi R_o}(\cos\theta_3 - \cos\theta_5)$ $b(R_o) = \frac{n_1}{\pi R_o}\left[\frac{R_o}{2}(\cos 2\theta_3 - \cos 2\theta_5) - 2a(\sin\theta_3 + \sin\theta_5)\right]\frac{2D}{\pi R_o}(\sin\theta_3 - \sin\theta_5)$
14		$g(R_o) = \frac{2n_1}{\pi}\left(\frac{\pi}{4} + \frac{\theta_3}{2} + \frac{\sin 2\theta_3}{4}\right)$ $b(R_o) = \frac{-2n_1}{\pi}\cos^2\theta_3$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

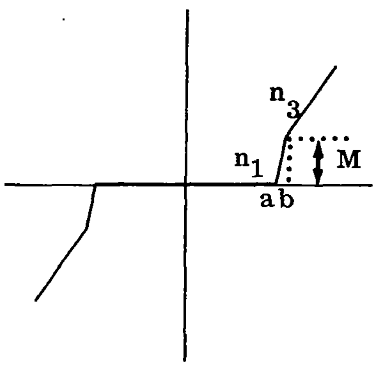
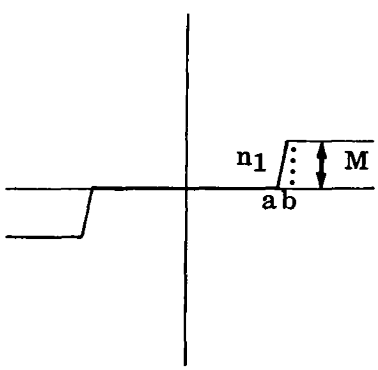
Identifying classification	Characteristic	Describing function
15		$g(R_o) = \frac{n_1}{\pi R_o}\Big[2 R_o(-\theta_1 + \theta_2) + R_o (\sin 2\theta_1 - \sin 2\theta_2 + 4a(\cos\theta_2 - \cos\theta_1))\Big] + \frac{n_3}{\pi R_o}\Big[R_o(\pi - 2\theta_2) + R_o \sin 2\theta_2 - 4k_3 \cos\theta_2\Big]$ $b(R_o) = 0$
16		$g(R_o) = \frac{n_1}{\pi R_o}\Big[2 R_o(-\theta_1 + \theta_2) + R_o(\sin 2\theta_1 - \sin 2\theta_2) + 4a(\cos\theta_2 - \cos\theta_1)\Big] + \frac{4M}{\pi R_o}\cos\theta_2$ $b(R_o) = 0$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

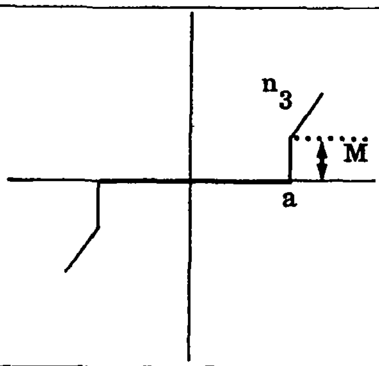
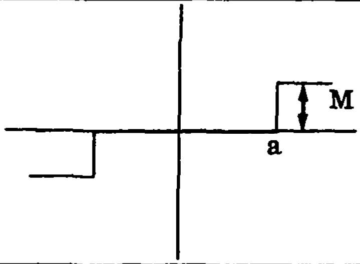
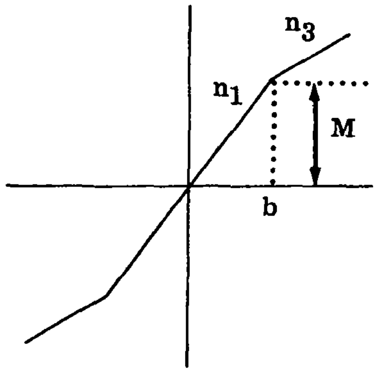
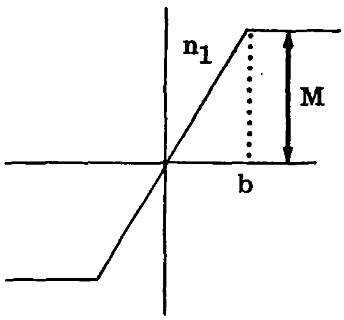
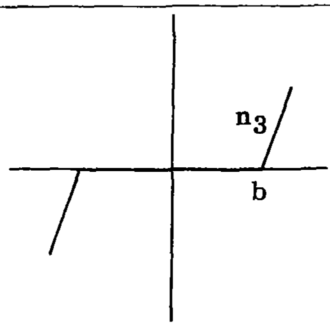
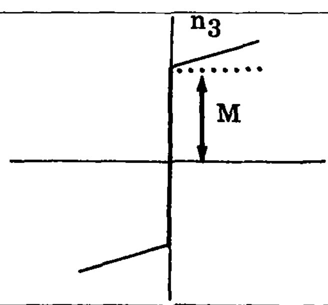
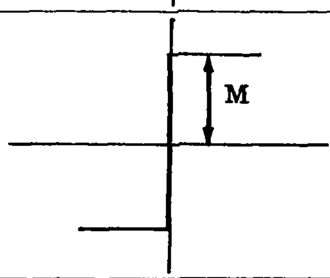
Identifying classification	Characteristic	Describing function
17		$g(R_o) = \frac{n_3}{\pi R_o}\left[R_o(\pi - 2\theta_1) + R_o \sin 2\theta_1 - 4k_3 \cos\theta_1\right]$ $b(R_o) = 0$
18		$g(R_o) = \frac{4M}{\pi R_o} \cos\theta_1$ $b(R_o) = 0$
19		$g(R_o) = \frac{n_1}{\pi}(2\theta_2 - \sin 2\theta_2) + \frac{n_3}{\pi R_o}\left[R_o(\pi - 2\theta_2) + R_o \sin 2\theta_2 - 4k_3 \cos\theta_2\right]$ $b(R_o) = 0$

Table 1. The Describing Functions for Twenty-three Common Nonlinear Elements, Contd

Nonlinear devices with memory

<table>
<tr><th>Identifying classification</th><th>Characteristic</th><th>Describing function</th></tr>
<tr><td>20</td><td>
</td><td>$$g(R_o) = \frac{n_1}{\pi}\left(2\theta_2 - \sin 2\theta_2\right) + \frac{4M}{\pi R_o}\cos\theta_2$$
$$b(R_o) = 0$$</td></tr>
<tr><td>21</td><td>
</td><td>$$g(R_o) = \frac{n_3}{\pi R_o}\left[R_o(\pi - 2\theta_2) + R_o \sin 2\theta_2 - 4b\cos\theta_2\right]$$
$$b(R_o) = 0$$</td></tr>
<tr><td>22</td><td>
</td><td>$$g(R_o) = n_3 + \frac{4M}{\pi R_o}$$
$$b(R_o) = 0$$</td></tr>
<tr><td>23</td><td>
</td><td>$$g(R_o) = \frac{4M}{\pi R_o}$$
$$b(R_o) = 0$$</td></tr>
</table>

Remark: All of the describing functions encountered thus far are independent of frequency. Occasionally, a situation arises in which the describing function is dependent on input frequency. This may occur, for example, when there is a velocity saturation. The analysis in this case is extremely complex and the use of a computer is mandatory. Various problems of this type are discussed by Kochenburger[11].

3.2.3 Application to Stability Analysis

The describing functions, as developed in the previous section, may be used to study the existence of self-sustained oscillations (limit cycles) in feedback control systems. It is important to bear in mind the major assumptions on which this analysis is based; namely,

a. The system is autonomous (i.e., unforced and time invariant).

b. The nonlinearity is frequency independent.

c. The linear transfer function of the system contains sufficient low pass filtering to warrant excluding from consideration all harmonics of the nonlinear element output other than the fundamental.

For purposes of analysis, the system under consideration may be depicted as shown in Fig. 33 with $r = 0$. We assume that the input to the nonlinearity may be represented as

$$x = R_o \sin \omega t$$

We now seek to determine the conditions which are sufficient to ensure that this oscillation be maintained indefinitely. If the system under discussion were linear ($G_D = 1$ in Fig. 33), then a sufficient condition for the existence of sustained oscillations would be

$$GH = -\frac{1}{K}$$

In the present case, this condition becomes

$$KGH = -\frac{1}{G_D(R_o)}$$

Furthermore, in the linear case, the simplified Nyquist criterion* states that if the critical point, $-1/K$ is to the left of the $G(j\omega)\, H(j\omega)$ locus when the latter

*The simplified Nyquist criterion is valid when all the poles of the open-loop transfer function lie in the left-hand plane.

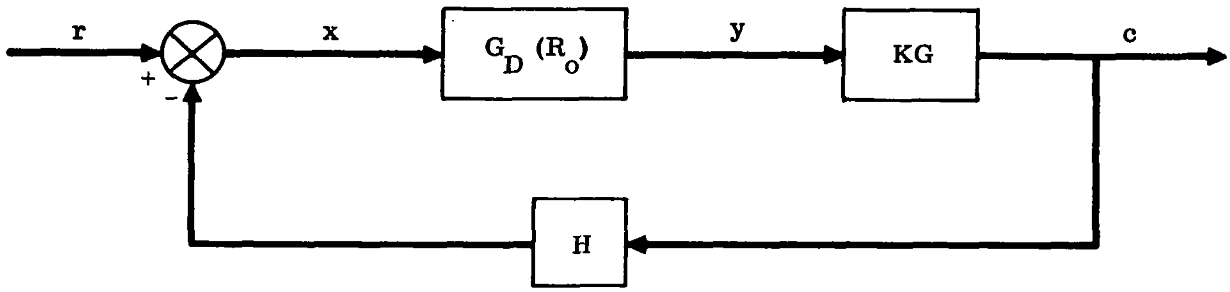


Figure 33. Use of Describing Function in Stability Analysis

is traced out in the direction of increasing frequencies, then the system is stable. In the nonlinear case, when the describing function is used to represent the nonlinearity, the above statement remains valid if $-\frac{1}{K}$ is replaced by $-\frac{1}{G_D(R_o)}$.

To examine the implication of these concepts, we plot the KGH and $-\frac{1}{G_D(R_o)}$ curves in the Nyquist plane, a typical case being as shown in Fig. 34. Here, the arrows on KGH and $-\frac{1}{G_D(R_o)}$ indicate the direction of increasing ω and R_o respectively. For a sufficiently low open-loop gain (the K_1 GH curve) there is no intersection of the curves, and no limit cycles exist. For a higher open-loop gain (the K_2 GH curve) a point of intersection is located at p_2, and a limit cycle is indicated having an amplitude and frequency corresponding to the values of R_o and ω at the point p_2. It remains to determine whether this limit cycle is stable or unstable. We note that when the value of R_o is such that $-\frac{1}{G_D(R_o)}$ is to the right of p_2, then instability is indicated by the simplified Nyquist criterion, causing R_o to increase and thereby moving the point $-\frac{1}{G_D(R_o)}$ to the left. When R_o is such that the point $-\frac{1}{G_D(R_o)}$ is to the left of p_2, then the Nyquist criterion indicates stability. Consequently, R_o decreases, moving the point $-\frac{1}{G_D(R_o)}$ to the right. It is obvious therefore that p_2 is a stable point of equilibrium and that p_2 is a stable limit cycle.

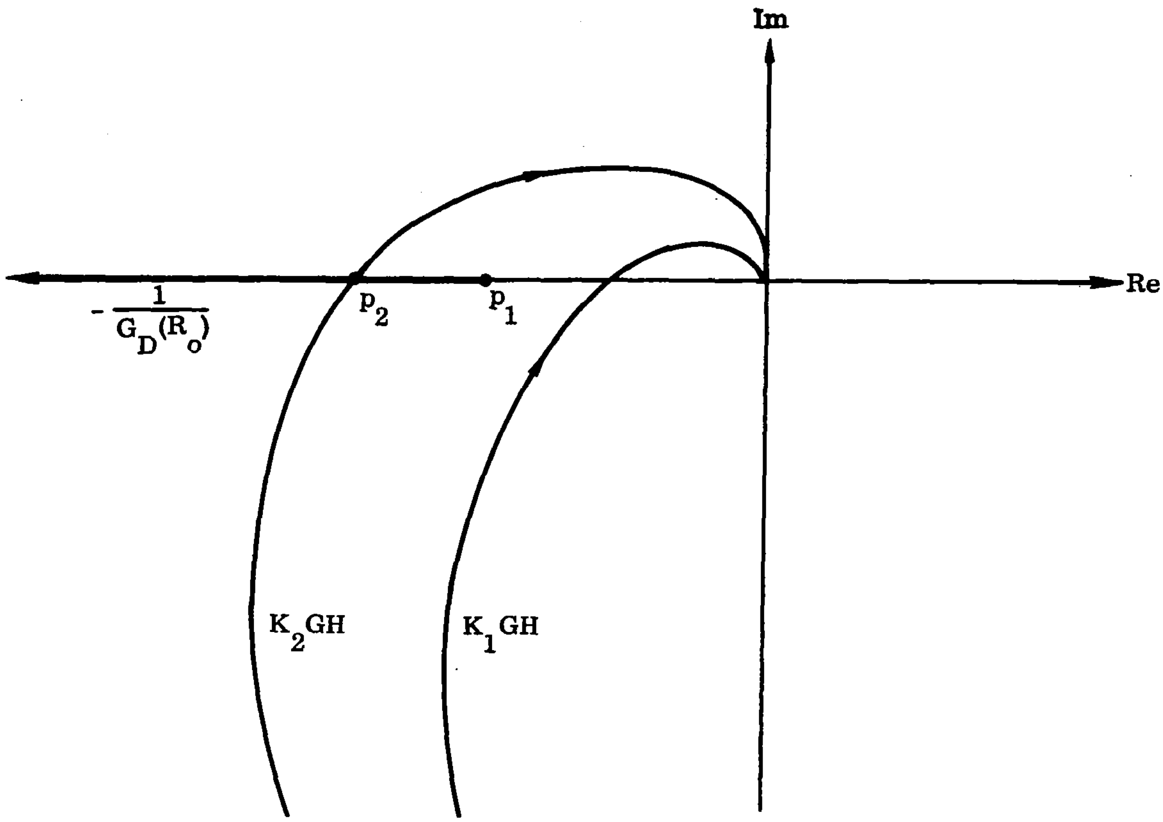


Figure 34. Nyquist Curves and Describing Function

The amplitude and frequency of this limit cycle is independent of the initial conditions, a phenomenon not exhibited by linear systems. When there is a phase lag associated with the describing function, the $-\frac{1}{G_D(R_o)}$ curve may appear as in Fig. 35. This particular representation is typical of a relay having hysteresis and dead zone. It may be shown that the Nyquist criterion still applies even though the critical point is not on the negative real axis. Thus, if R_o is such that the operating point is to the right of p_2, the system is unstable, R_o increases, and the operating point moves along the $-\frac{1}{G_D(R_o)}$ curve towards p_2. Reasoning as before, we find that p_2 represents a point of stable equilibrium; i.e., a stable limit cycle.

Fig. 36 indicates a situation with increased dead zone. The result is that there is no intersection of the curves and no limit cycle can exist.

Another possibility is shown in Fig. 37 where the curves intersect at two points. To determine the stable limit cycle we may reason as follows. If R_o is such that the operating point on $-\frac{1}{G_D(R_o)}$ is located between p_1 and p_2, the system is stable, R_o decreases, and the operating point moves to p_1. The motion is stable and there is no limit cycle. If the operating point on the $-\frac{1}{G_D(R_o)}$ curve is between p_2 and p_3, the motion is unstable, R_o increases and the operating point moves to p_3. Also, if the operating point is to the left of p_3, the motion is stable, and the decreasing R_o

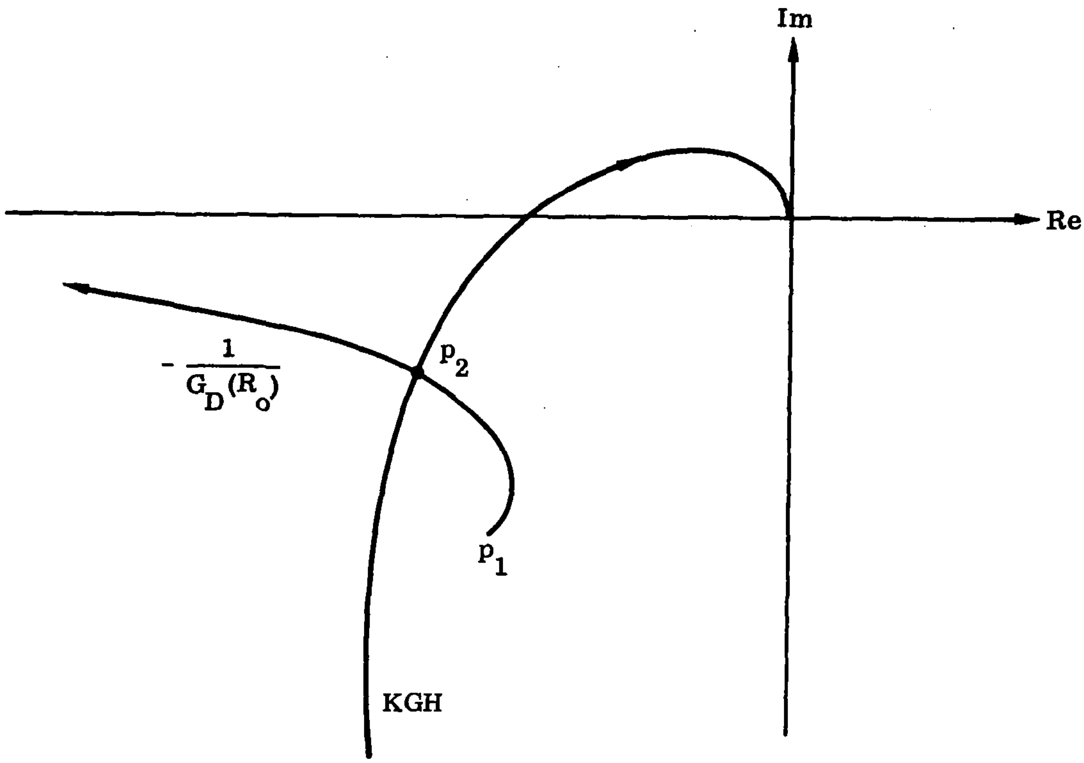


Figure 35. Nyquist Curve and Describing Function with Memory -- Limit Cycle Shown

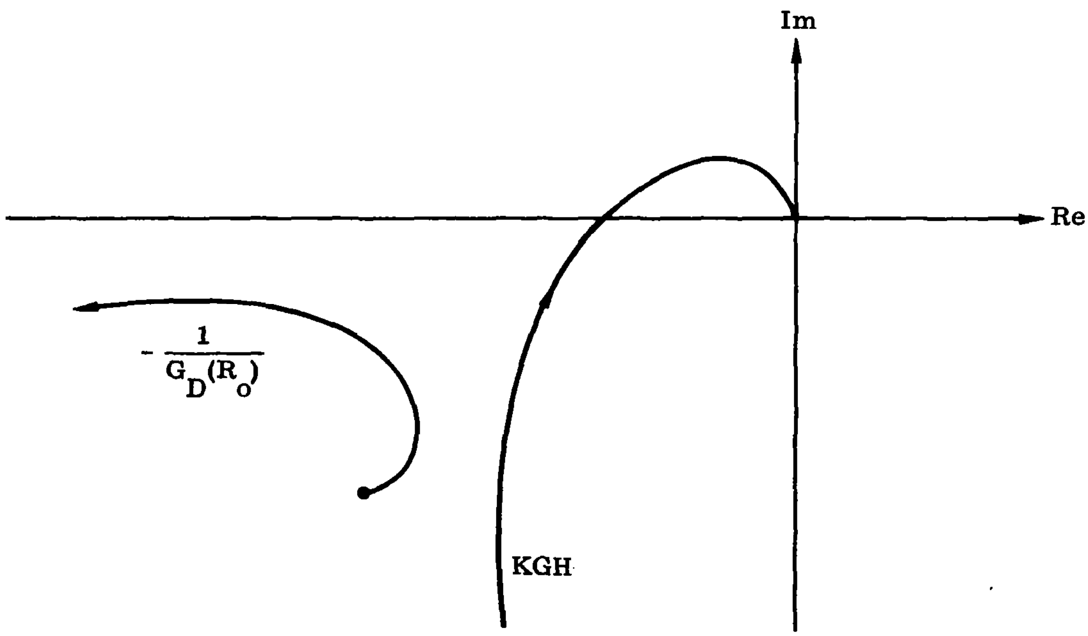


Figure 36. Nyquist Curve and Describing Function with Memory -- No Limit Cycle

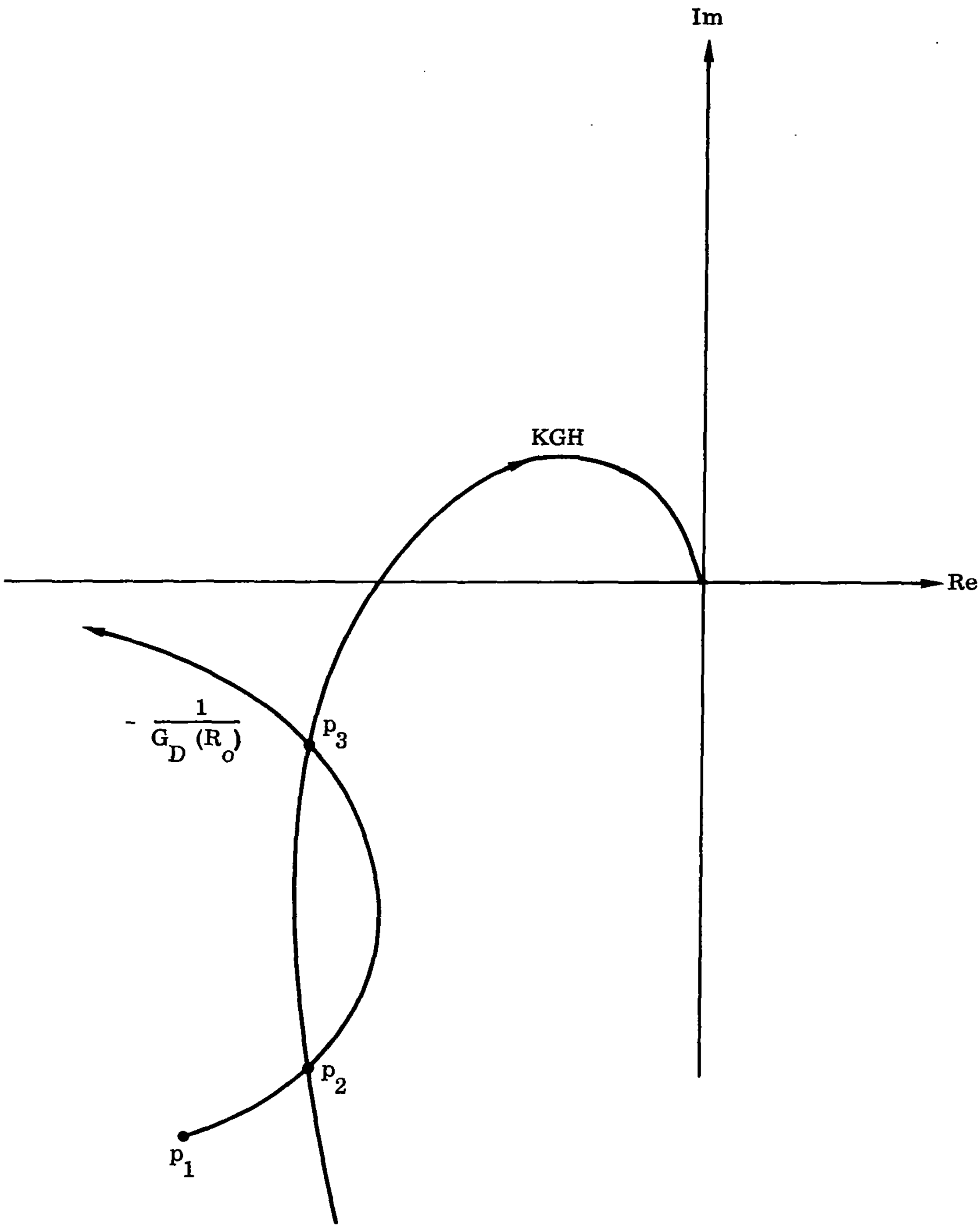


Figure 37. Nyquist Curve and Describing Function with Memory -- Two Limit Cycles Shown

causes the operating point to move toward p_3. Thus, p_3 represents a stable limit cycle and p_2 is an unstable limit cycle. Summarizing, we find that any disturbance that causes an R_o whose operating point on the $-\frac{1}{G_D(R_o)}$ curve is to the left of p_2 will result in a motion which damps out completely, while larger amplitudes of R_o will result in a limit cycle of frequency and amplitude determined by point p_3.

3.2.4 Multiple Nonlinearities

The analyses of the previous sections have been based on the assumption that there is only one nonlinearity in the system. For multiple nonlinearities, various extensions of the describing function technique have been described in the literature.(13) - (15) Perhaps the most straightforward approach is to develop a composite describing function when the nonlinear elements are in series. This problem has been studied by Gronner.(15) If the nonlinearities are separated by linear dynamic elements, the analysis is much more difficult. Even for relatively simple systems, some sort of computer assist is required.

While a detailed computational approach is tedious, the broad outlines of the method of analysis are relatively simple and afford a degree of insight to the results generated by computer. We will here describe a technique due to Gran and Rimer(12) which is applicable to multiple nonlinearities which can be put in the form shown in Figs. 38 or 39. Certain restrictions must be imposed on the linear elements, $G_i(s)$, for the describing function method to be applicable. In Fig. 38, there is no restriction as long as the input signal, x, is sinusoidal, which will be the case if the system is low pass. For the form of Fig. (39), however, the output of the linear dynamic elements must be sinusoidal. Therefore, all the $G_i(s)$ must be low pass.

The method to be described is valid for various modifications or combinations of Figs. 38 and 39. The essential condition that must be satisfied is that the input to the nonlinear elements be sinusoidal; this requirement serves to indicate which parts of the system must be low pass.

Consider now the system of Fig. 38. The input, x, is assumed to be of the form

$$x = R_o \sin \omega t$$

Consequently, the input to the i^{th} nonlinearity is

$$R_o \, |G_i(j\omega)| \sin (\omega t + \angle G_i(j\omega))$$

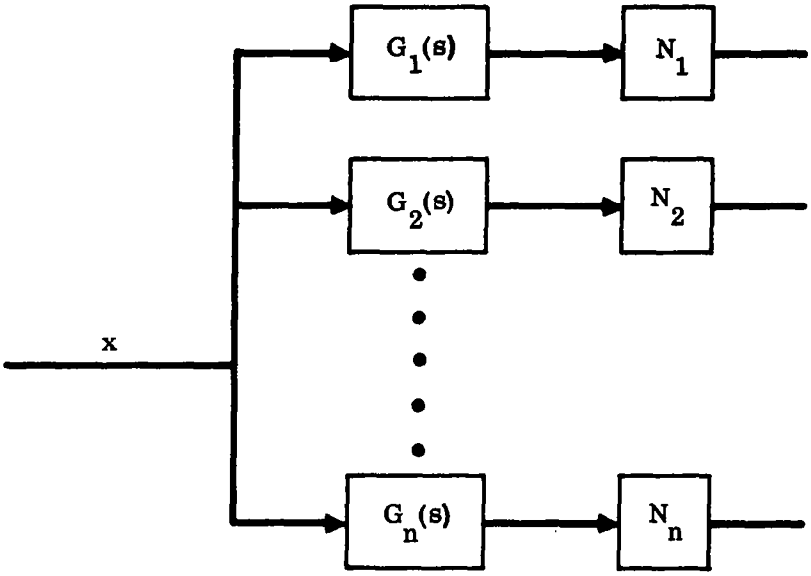


Figure 38. Nonlinear Elements in Parallel

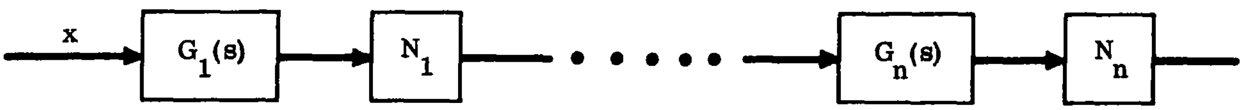


Figure 39. Nonlinear Elements in Series

The describing function of the i^{th} nonlinearity, denoted by $G_D^{(i)}(R_o, \omega)$, is therefore a function of the amplitude, R_o, and the input frequency, ω. If these describing functions are used to replace the nonlinearities in the system block diagram, the result is a linear system.

In Fig. 39, the input to the first nonlinearity is

$$R_o \, |G_1 (j\omega)| \sin (\omega t + \angle G_1 (j\omega))$$

The input to the second nonlinearity becomes

$$R_o \, |G_1 (j\omega)| \; |G_2 (j\omega)| \; G_D^{(1)} (R_o, \omega) \sin (\omega t + \angle G_1 + \angle G_2)$$

and this is used to determine the describing function for the second nonlinearity, $G_D^{(2)}(R_o, \omega)$. This procedure is continued until all the describing functions are formulated. Although these are in general quite complex, they depend only on R_o and ω. For either form, a set of values of R_o and ω will determine a linear system whose operating gain and pole-zero locations are determined for these values of R_o and ω.

The characteristic equation for the entire system is now written as

$$1 + K_o (R_o, \omega) \frac{\prod_{i=1}^{N} [s + Z_i (R_o, \omega)]}{\prod_{i=1}^{M} [s + P_i (R_o, \omega]} = 0 \tag{72}$$

For a specified value of R_o and ω, a pole-zero configuration is determined along with an operating gain, K_o. A limit cycle is indicated at that point on the root locus where the latter intersects the imaginary axis. However, a change in R_o (or ω) also changes the pole-zero configuration as well as the operating gain. Therefore, in general, a separate root locus is required for every value of R_o and ω. The general procedure is as follows.

a. Select a value of R_o and ω (say $R_o^{(1)}$ and ω_1). Determine the root locus and operating gain, $K_o^{(1)}$.

b. Select a new value of R_o (say $R_o^{(2)}$), and with the same value of ω_1. A new pole-zero configuration and a new root locus result along with a new value for operating gain, $K_o^{(2)}$.

c. On a separate diagram, connect all the operating points, $K_o^{(i)}$, resulting in an "input-dependent locus." The limit cycle amplitude and frequency are determined by the intersection of this locus with the $j\omega$ axis. If the frequency at crossover and the original assumed frequency, ω_1, do not correspond, then a new input-dependent locus" for a different ω, must be determined and the process continued until the assumed frequency and crossover frequency correspond. This is the predicted limit cycle.

The above procedure simplifies materially for the case of Fig. 38 since the pole-zero locations and operating gains do not depend on frequency.

In any event, the "input-dependent locus" generated by this method may be interpreted as follows:

a. If the locus remains in the right-hand plane, the system is unstable; if it remains in the left-hand plane, the system is stable.

b. If an increasing R_0 causes the locus to cross the imaginary axis from the right-hand to the left-hand plane, a stable limit cycle exists.

c. If increasing R_0 causes the locus to cross the imaginary axis from the left-hand to the right-hand plane, there is an unstable limit cycle.

The proof of the above assertions follows the conventional arguments for root locus stability. As an example of the above approach, we consider the following.

Example 6: The block diagram of the system to be investigated is shown in Fig. 40. This may be rearranged to the form shown in Fig. 41, which is of the parallel form of Fig. 38. Now by putting $x = R_0 \sin \omega t$, the open-loop transfer function becomes

$$\frac{K_o(R_o)}{(s+1)^2 \left(s + \frac{1}{\tau}\right)}$$

where

$$K_o(R_o) = \frac{100\, G_D^{(1)}(R_o)}{10\, G_D^{(2)}(R_o) + 1}$$

$$\frac{1}{\tau} = \frac{100\, G_D^{(2)}(R_o) + 1}{10\, G_D^{(2)}(R_o) + 1}$$

$$G_D^{(1)}(R_o) = \frac{2n_1}{\pi}\left[\sin^{-1}\frac{b}{R_o} - \frac{b}{R_o^2}\sqrt{R_o^2 - b^2}\right] + \frac{4M_1}{\pi R_o^2}\sqrt{R_o^2 - b^2}$$

$$G_D^{(2)}(R_o) = \frac{4M_2}{\pi R_o}$$

n_1 ≡ slope of saturation nonlinearity.

In this example

$b = 1$

$M_1 = 1000$

$M_2 = 10$

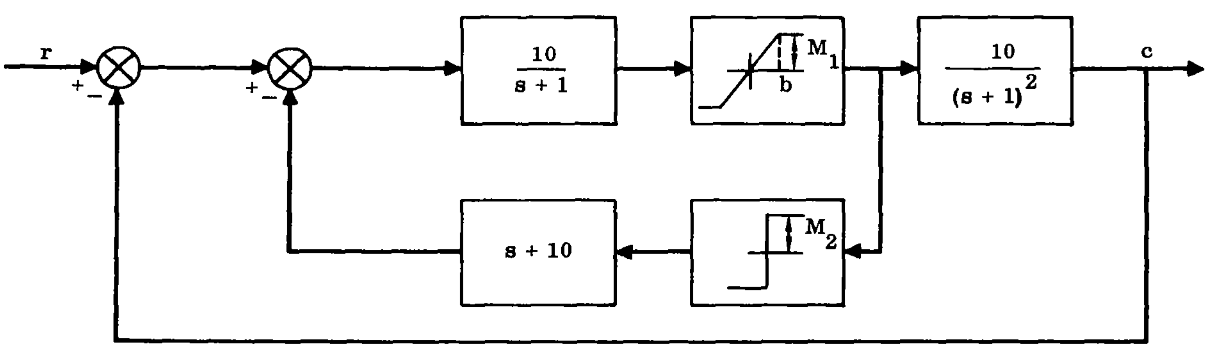


Figure 40. Control Loop for Example 6.

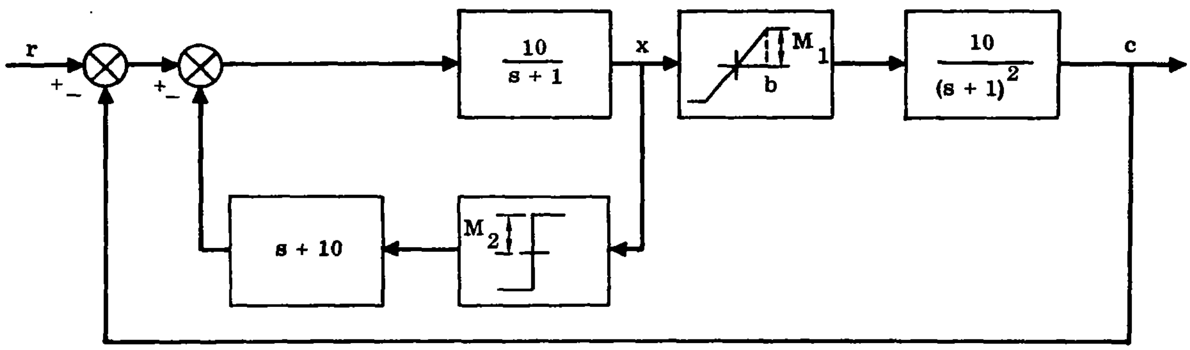


Figure 41. Equivalent Control Loop for Example 6

A typical root locus for a particular value of R_o is shown in Fig. 42, while Fig. 43 depicts the "input-dependent locus" for this problem. The pole locations for specified values of R_o are summarized in the following table.

Point on Locus	R_o	Pole Location	K_o
1	0.1	-10.00	78
2	0.4	-10.00	250
3	0.5	-10.00	396
4	1.0	-9.96	781
5	4.0	-9.70	960
6	100.0	-6.05	879
7	1000.0	-2.01	177

For $R_o = 0$, the point on the input-dependent locus corresponds to the double pole at -1, and for $R_o \rightarrow \infty$, the locus returns to this same point.

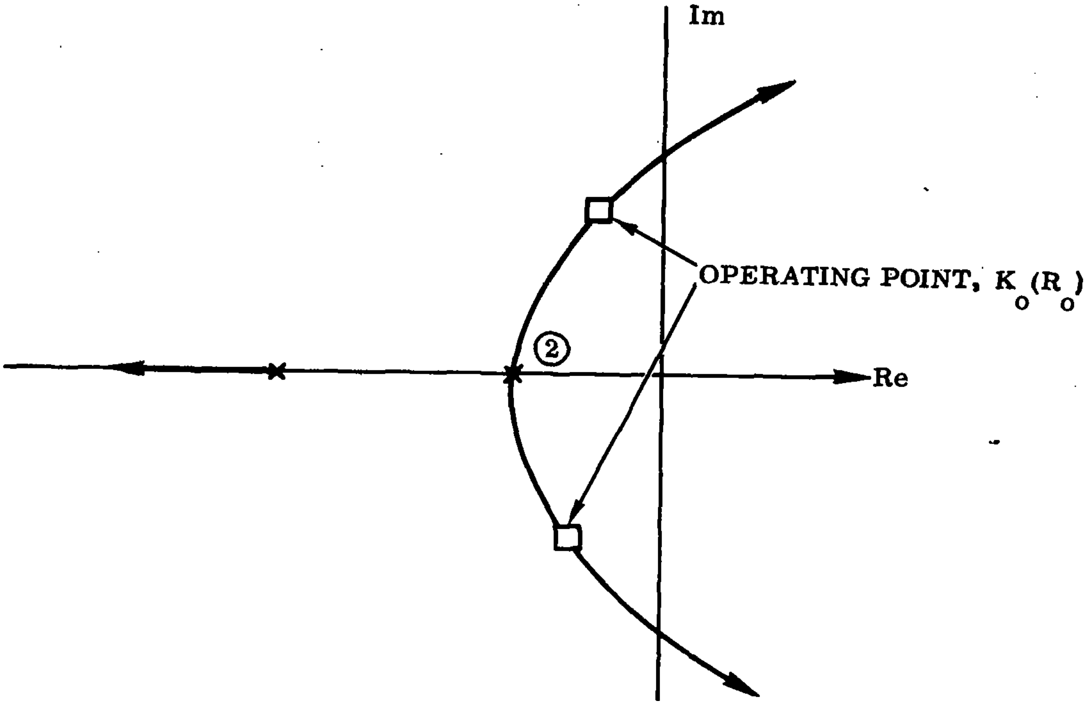


Figure 42. Root Locus for Example 6

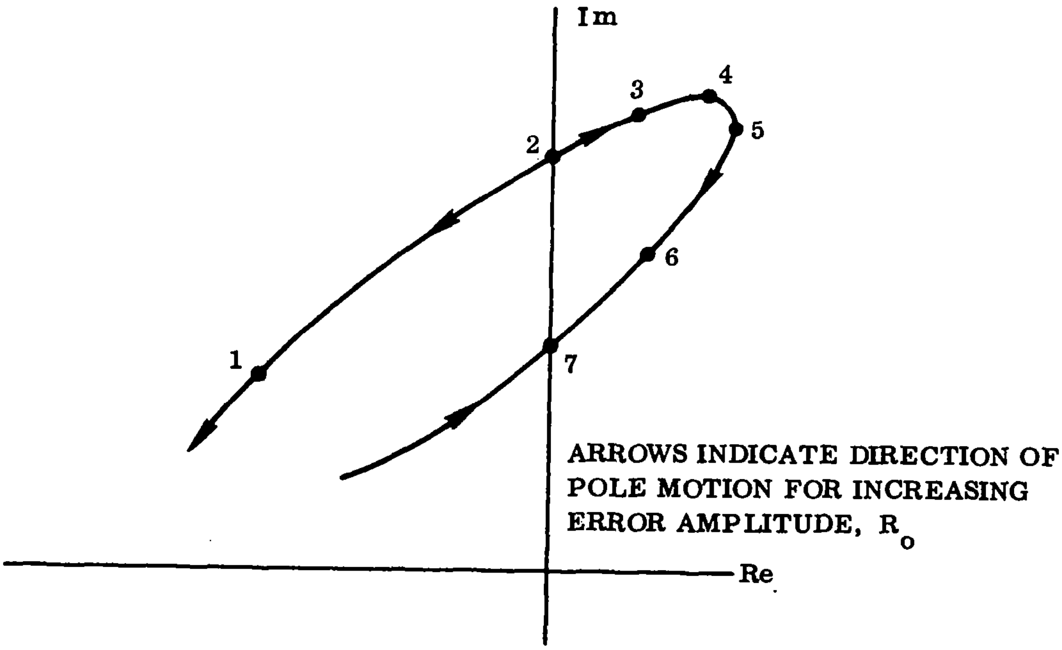


Figure 43. Gain Locus for Example 6

In accordance with the criteria listed above, the point 2 in Fig. 43 corresponds to an unstable limit cycle while point 7 corresponds to a stable limit cycle. The latter occurs at the values, $R_o = 1000$ and $\omega = 3.6$ rad/sec.

3.2.5 Dual Input Describing Function (DIDF)

The method is based on consideration of two sinusoids being applied at the input of a nonlinear element. In certain cases a high frequency low amplitude signal (dither) is introduced externally to modify the characteristics of the nonlinearity. This is the so-called "dynamic lubrication" method, which often has the effect of linearizing the nonlinear element and tends to stabilize the system.

In the present case, however, the high frequency signal is taken to be the system limit cycle while the slowly varying signal is derived from input commands to the system. The concept of a DIDF was first introduced by West(42) and is discussed in a standard text by Gibson(43). While useful in certain cases, this approach is extremely tedious and lengthy. An extension of the method by Gelbe and Vander Velde(16) affords a crucial simplification in that a limit cycling system is treated as a control system whose signal transmission properties can be readily derived under certain mild restrictions on the input command signal. The discussion will be limited to feedback control systems of the form shown in Fig. 44.

If a system which exhibits a limit cycle in the unforced state is subjected to a slowly varying input, the output follows this input, on the average, to within some following error. A typical situation is shown in Fig. 45. Over any limit cycle period, the system error may be approximately modeled as a sinusoid plus d-c bias. The sinusoid is associated with the limit cycle and the d-c bias with the error. This suggests that the input to the nonlinear element be taken as

$$x = R_1 + R_o \sin \omega_o t \tag{73}$$

Making the same assumptions as for conventional describing functions, the output from the nonlinear element becomes

$$y = R_1' + R_o' \sin (\omega_o t + \varphi_1) \tag{74}$$

In other words, the output from the nonlinearity is a different d-c bias plus a sinusoid of different amplitude and with some phase angle, φ_1, with respect to the input. The quantity, ω_o, represents the limit cycle frequency. We now define the following:

$$\text{Limit Cycle DIDF} \equiv N_o (R_o, R_1) = \frac{R_o'}{R_o} e^{j\varphi_1} \tag{75}$$

$$\text{Signal DIDF} \equiv N_s (R_o, R_1) = \frac{R_1'}{R_1} \tag{76}$$

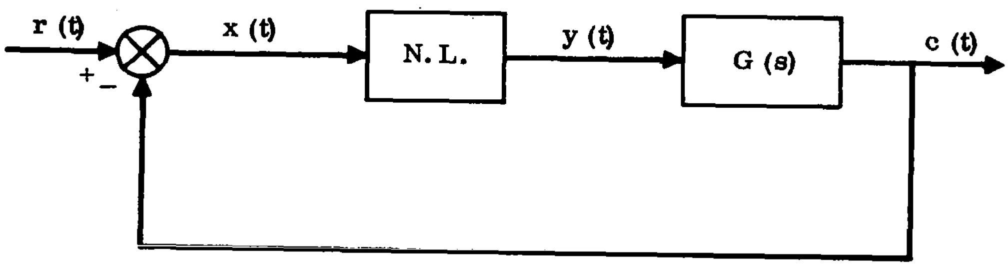


Figure 44. Control Loop for DIDF Analysis

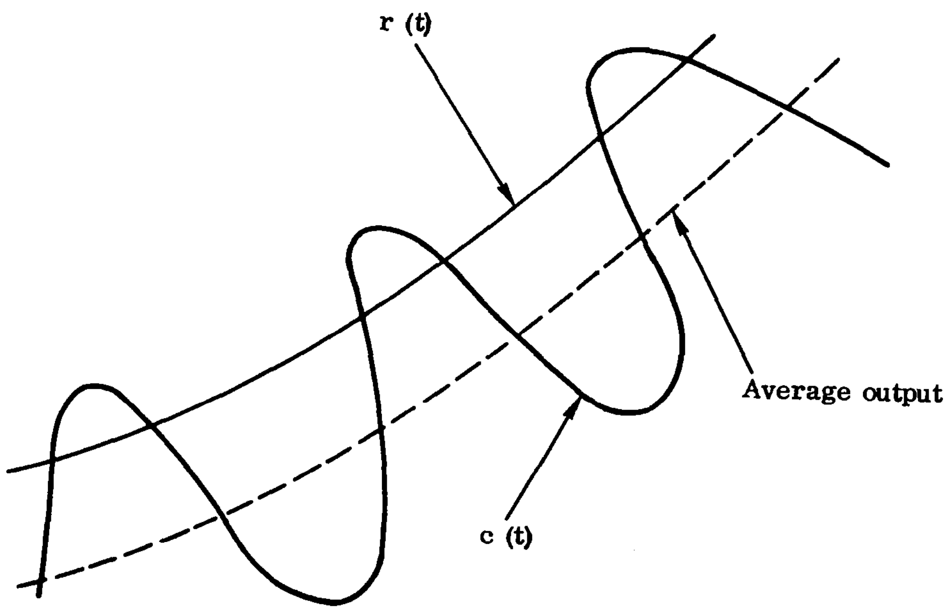


Figure 45. Input and Output Waveforms

The nonlinearity in a limit cycling control system is therefore characterized by two DIDF's. The first of these, the limit cycle DIDF is used to predict the system limit cycle in a manner identical to that in conventional describing function analysis; i.e., the limit cycle is obtained as the solution to the equation

$$N_o\,(R_o,\,R_1)\;G(j\,\omega) = -1$$

as described in Section 3.2.3.

The signal DIDF will be considered as a linear gain for some range of (R_1/R_o), and will therefore lead to a linear system description from the point of view of command inputs. This situation is depicted in Fig. 46, and represents the approximate mathematical model for the study of the input-output dynamics of the given nonlinear system.

It is clear that the equivalent representation is valid whenever the DIDF formulations are themselves valid. This means that in addition to the usual conditions imposed on the conventional describing function, two additional requirements must be satisfied. These are:

$$\begin{aligned} \frac{R_1}{R_o} &< \frac{1}{3} \\ \frac{\omega_s}{\omega_o} &< \frac{1}{3} \end{aligned} \tag{77}$$

where ω_s is the frequency of the input signal, x(t), and ω_o is the limit cycle frequency. A thorough discussion of these conditions is contained in Gelb and Vander Velde(16).

The requirement imposed on the frequency ratio stems from the assumption that the error (between input and output) be approximated by a d-c bias over one limit cycle period. The requirement on amplitude ratio is necessary to ensure that higher harmonics in the DIDF approximation be negligible compared to the fundamental. Gelb and Vander Velde show that when conditions (77) are satisfied, the approximate DIDF is within 5% error referred to the untruncated expression for the DIDF.

The above ideas can be clarified by considering a particular case.

Example 7: The system to be investigated is shown in Fig. 47. We begin by deriving the relevant DIDF's for the relay. We use

$$x = R_1 + R_o \sin \omega_o t$$

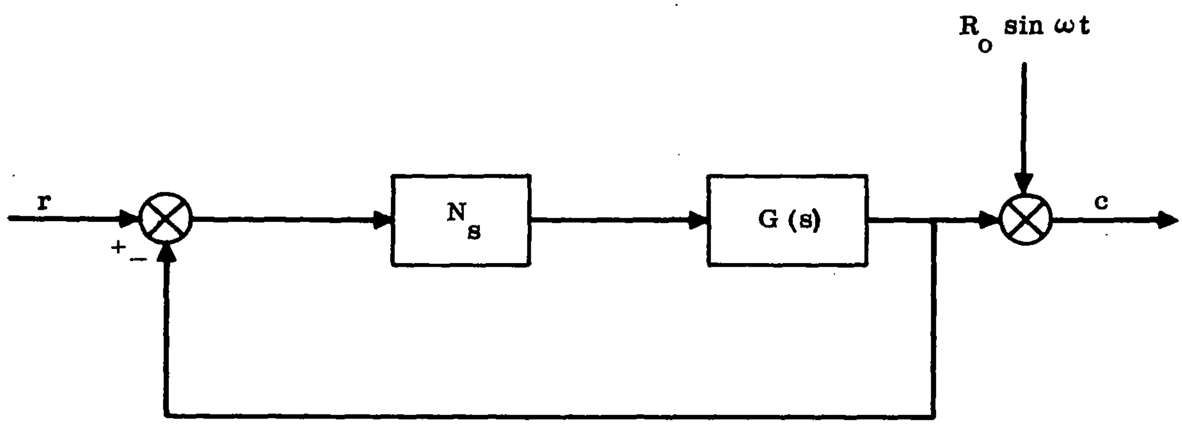


Figure 46. Equivalent Control Loop for DIDF Analysis

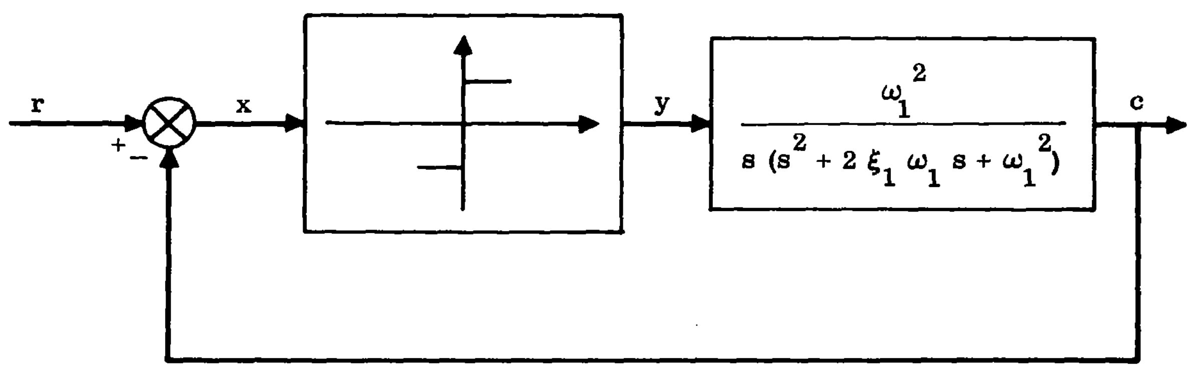


Figure 47. Control Loop for Example 7

From Eq. (75)

$$N_o = \frac{2}{\pi R_o}\left[\int_{-\frac{\pi}{2}}^{-\theta_1}(-D)\sin\theta\, d\theta + \int_{-\theta_1}^{\frac{\pi}{2}} D\sin\theta\, d\theta\right]$$

$$= \frac{4D}{\pi R_o}\sqrt{1-\left(\frac{R_1}{R_o}\right)^2}$$

where $\theta \equiv \omega_o r$. Also

$$N_s = \frac{1}{2\pi R_1}\left[\int_{o}^{\pi+\theta_1} D\, d\theta - \int_{\pi+\theta_1}^{2\pi-\theta_1} D\, d\theta + \int_{2\pi-\theta_1}^{2\pi} D\, d\theta\right]$$

$$= \frac{2D}{\pi R_1}\theta_1$$

where $\theta_1 = \sin^{-1}\left(\frac{R_1}{R_o}\right)$. Note that for R_1 "small"

$$N_o \approx \frac{4D}{\pi R_o}$$

which is precisely the conventional describing function for a relay, and

$$N_s \approx \frac{2D}{\pi R_o}$$

The aim of the present analysis is to determine the conditions under which a representation of the form of Fig. 46 is valid. For in that case, conventional linear techniques may be employed to determine the input-output system dynamics. It is not difficult to show that in order to satisfy the relations in Eq. (77), it is necessary that the dominant closed-loop poles of G(s) in Fig. 46 lie within a circle of radius $\frac{1}{3}\omega_o$ (with center at the origin) in the complex plane.

The limit cycle frequency, ω_0, and amplitude, R_0, may be determined in the usual manner for the system of Fig. 47. The results are displayed in the root locus of Fig. 48. For this simple case, it is more convenient to work with the equation of the root locus which is

$$3\sigma^2 - \omega^2 + 4\,\xi_1\,\omega_1\,\sigma + \omega_1^2 = 0$$

$$s = \sigma + j\,\omega$$

This curve crosses the imaginary axis ($\sigma = 0$) when $\omega = \omega_1$ which means that the limit cycle frequency, $\omega_0 = \omega_1$.

Now by substituting $s = j\omega_1$ in

$$1 + N_o\,G(S) = 0$$

we find

$$N_o = 2\,\xi_1\,\omega_1$$

But since $N_o \approx \frac{4D}{\pi R_o}$, the limit cycle amplitude is determined as

$$R_o \approx \frac{2\,D}{\pi \xi_1\,\omega_1}$$

The value of N_s in Fig. 46 is, in the present case,

$$N_s \approx \frac{N_o}{2} = \xi_1\,\omega_1$$

In order to determine the value of the closed-loop pole on the real axis, we put

$$s = \sigma$$

$$N_s = \xi_1\,\omega_1$$

in the system characteristic equation

$$1 + N_s\,G(s) = 0$$

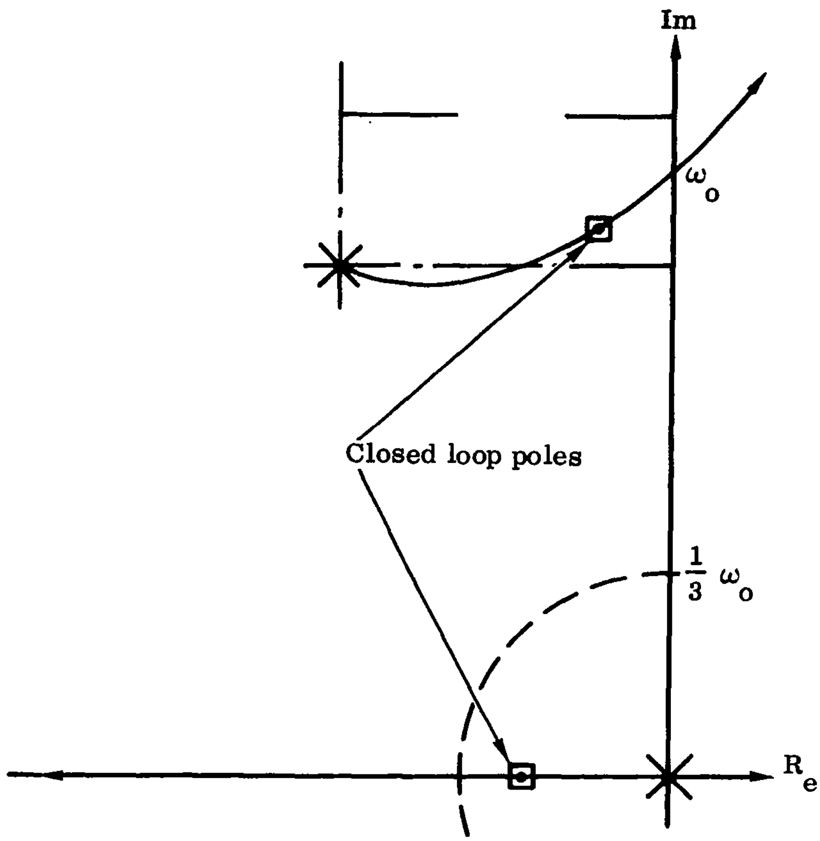


Figure 48. Root Locus for Example 7

This yields $\sigma \approx -\xi_1 \omega_1$ if $\xi_1 < 0.1$. The closed-loop poles are indicated in Fig. 48. It is seen that the dominant closed-loop pole is within the circle of radius $\frac{1}{3}\omega_o$, indicating that the equivalent representation of Fig. 46 is valid. A typical response to a step input is shown in Fig. 49 for the actual system and for its equivalent linear representation. The agreement between the dominant modes in the two cases is seen to be remarkably good.

3.2.5.1 DIDF for Typical Nonlinearities

The DIDF for some of the more common types of nonlinearities are listed below.

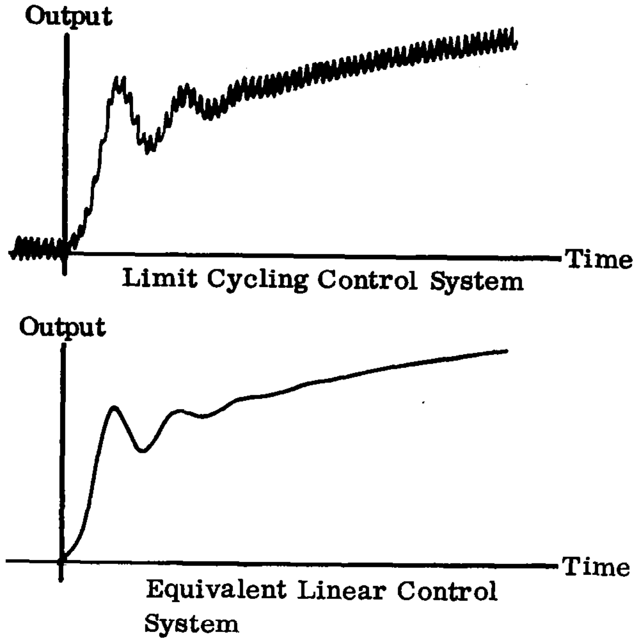


Figure 49. Response to Step Input for System of Example 7

a. Relay with Dead Zone (Fig. 50a)

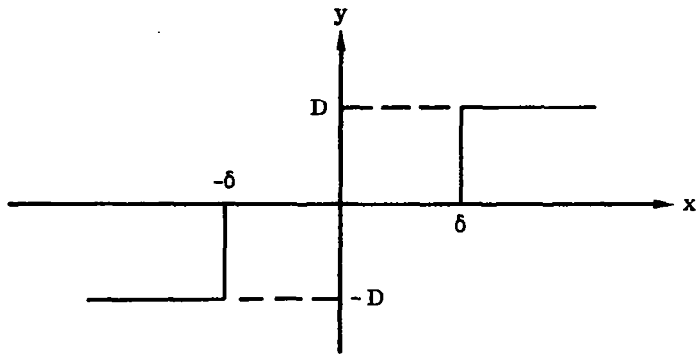
$$N_o = \frac{4D}{\pi R_o}\left[1 - \frac{1}{2}\left(\frac{\delta}{R_o}\right)^2 - \frac{1}{2}\left(\frac{R_1}{R_o}\right)^2\right]$$

$$N_s = \frac{2D}{\pi R_o}\left[1 + \frac{1}{2}\left(\frac{\delta}{R_o}\right)^2 + \frac{1}{6}\left(\frac{R_1}{R_o}\right)^2\right]$$

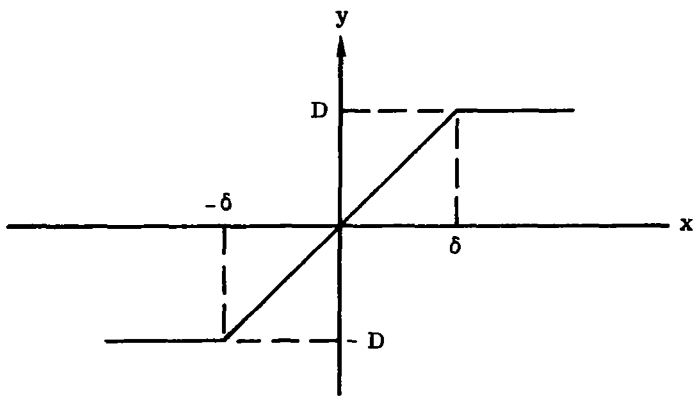
b. Saturating Element (Fig. 50b)

$$N_o = \frac{4D}{\pi R_o}\left[1 - \frac{1}{6}\left(\frac{\delta}{R_o}\right)^2 - \frac{1}{2}\left(\frac{R_1}{R_o}\right)^2\right]$$

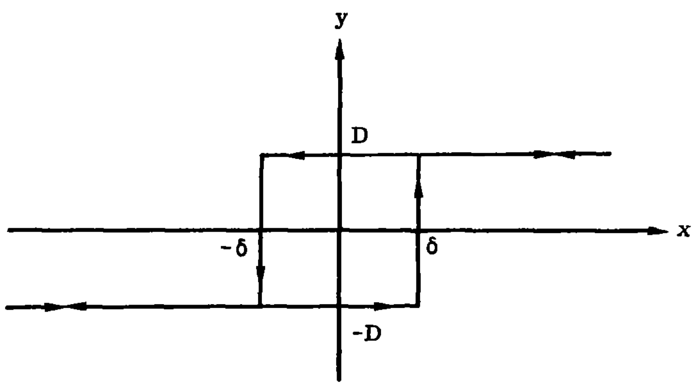
$$N_s = \frac{2D}{\pi R_o}\left[1 + \frac{1}{6}\left(\frac{\delta}{R_o}\right)^2 + \frac{1}{6}\left(\frac{R_1}{R_o}\right)^2\right]$$



a. Relay with Dead Zone



b. Saturation



c. Hysteresis

Figure 50. Typical Nonlinear Characteristics

c. Rectangular Hysteresis (Fig. 50c)

$$N_o = \frac{4D}{\pi R_o}\left[1 - \frac{1}{2}\left(\frac{R_1}{R_o}\right)^2\right] e^{-j\frac{\delta}{R_o}}$$

$$N_s = \frac{2D}{\pi R_o}\left[1 + \frac{1}{2}\left(\frac{\delta}{R_o}\right)^2 + \frac{1}{6}\left(\frac{R_1}{R_o}\right)^2\right]$$

d. Polynomial Type Nonlinearity. This nonlinearity is expressible in the form

$$y = C_n x^n + C_{n-1} x^{n-2} |x| + \dots\dots\dots + C_2 x |x| + C_1 x$$

There is no loss of generality in assuming that n is odd. We have

$$N_o = \frac{2}{\sqrt{\pi}} \sum_{k \text{ even}}^{n} \frac{C_k n!}{(n-k)!\,k!} R_o^{n-k-1} R_1^k \frac{\Gamma\left(\frac{n-k+2}{2}\right)}{\Gamma\left(\frac{n-k+3}{2}\right)}$$

$$N_s = \frac{2}{\sqrt{\pi}} \sum_{k \text{ odd}}^{n} \frac{C_k n!}{(n-k)!\,k!} R_o^{n-k} R_1^{k-1} \frac{\Gamma\left(\frac{n-k+1}{2}\right)}{\Gamma\left(\frac{n-k+2}{2}\right)}$$

where Γ ( ) denotes the gamma function.

3.3 LYAPUNOV'S DIRECT METHOD

The general theory of stability of motion as developed by A. Lyapunov[17] (also known as Lyapunov's Direct Method or the Second Method of Lyapunov) is now recognized as the most powerful theoretical technique available for the study of nonlinear control systems. Indeed the extreme generality of the theory also accounts for the frustrating difficulties experienced in seeking to apply the theory to systems of even moderate complexity. It is important to recognize that the Lyapunov approach provides merely a framework for analysis rather than a detailed computational algorithm. This is perhaps to be expected since nonlinear systems are best defined in a negative sense; that is, a nonlinear system is merely one which is not linear.

The Lyapunov method was virtually unknown in this country as little as ten years ago. However, in recent years, a prodigious technical literature has been developed which has served to stimulate interest in this subject among control engineers. References 18 through 21 are concerned with the basic theory and contain extensive bibliographies.

For purposes of the present monograph, the exposition will be limited to a careful and a precise statement of the principal theorems in language that is familiar to control engineers. While this will somewhat compromise the rigor of presentation, it will hopefully be compensated by the added degree of physical insight obtained. Extensive use will be made of illustrative examples and physical applications. Of the principal directions in which the theory has developed in recent years, only those which are potentially useful in control system analysis will be stressed. Few proofs will be given and little attention will be paid to mathematical niceties. These will be disposed of with reference to the pertinent literature. All important results will, however, be stated as precisely and completely as possible, with due regard for inherent limitations and conditions to be satisfied. In this way, it is felt, the material can be put in a form which is most useful for purposes of practical application.

3.3.1 Symbols and Definitions

The presentation of the material in the following sections is greatly facilitated by the use of matrix methods and a uniform terminology. It will be assumed that the reader is familiar with the basic operations involving vectors and matrices. Throughout Section 3.3, the following conventions will be adopted:

a. Capital letters will denote matrices.

b. Lower case English letters will denote vectors.

c. Lower case Greek letters will denote scalars.

d. Subscripted English and Greek lower case letters will denote scalars.

e. The following exceptions are noted.

    1. V, e, g, h, r, t will denote scalars.

    2. i, j, k, n, will denote integers.

Also

A^T	$\equiv$	transpose of matrix A
A^{-1}	$\equiv$	inverse of matrix A
$a \cdot b$	$\equiv$	scalar product of vectors
∇a	$\equiv$	gradient of a.
$(\dot{\ })$	$\equiv$	$^d/dt\ (\ )$

Various terms will be used repeatedly in the following discussions, and, for clarity, these are listed and defined below.

Definition 1: A system is said to be linear if it may be represented as

$$\dot{x} = Ax + f(t) \tag{78}$$

where A is a constant matrix. If the coefficients of the matrix, A, are functions of the independent variable, t, then the system is said to be linear time varying. Any dynamical system which cannot be expressed in the form (78) is called nonlinear.

Definition 2: The variables, $x_1, x_2, \ldots\ldots\ldots, x_n$, which are the components of the state vector, x, are referred to as the state variables since at any given instant of time, they represent the condition or state of the system.

Definition 3: The n-dimensional Euclidean space which has axes labeled $x_1, x_2, \ldots\ldots,$ x_n is called the state space. Each point in this state space represents a particular set of values that the state variables may assume. In other words, each point represents a particular state of the system. Given the initial state of the system ($t = 0$), then for $t > 0$ the state variables describe a trajectory in state space.

Definition 4: A nonlinear system, described in general by

$$\dot{x} = g(x, t) + \psi(t) \tag{79}$$

and the linear system, given by Eq.(78)are said to be forced if f and ψ are not zero.

Definition 5: The systems described by Eqs. (78) and (79) are said to be free if f and ψ are zero.

Definition 6: The systems described by Eqs. (78) and (79) are said to be stationary if the elements of A and g are not functions of the independent variable, t; otherwise, they are called nonstationary.

Definition 7: A system that is both free and stationary is called autonomous.

Definition 8: Associated with every vector, x, is a scalar called the norm which is written as $\|x\|$. A norm satisfies the following relations:

1. $\|x\| > 0$ for all $x \neq 0$
2. $\|x\| = 0$ if $x = 0$
3. $\|x + y\| \leq \|x\| + \|y\|$
4. $\|\alpha x\| = |\alpha| \cdot \|x\|$, α = any scalar

The following are some commonly used norms.

1. $\|x\| = (x^T x)^{1/2} = \sqrt{\sum_i x_i^2}$

2. $\|x\| = \sum_i |x_i|$

3. $\|x\| = \max_i |x_i|$

Only the first of these, which is the standard concept of length of a vector, will be used in this monograph.

Definition 9: The scalar defined by

$$\mu = x^T Bx$$

is called a quadratic form in the state variables $x_1, x_2, \ldots\ldots, x_n$. The $n \times n$ matrix, B, is symmetric.

Definition 10: The quadratic form, μ, will be called positive definite if, for all $x \neq 0$ and such that $\mu = 0$ for $x = 0$,

$$\mu > 0$$

and negative definite if

$$\mu < 0$$

Furthermore, μ is said to be positive semidefinite if

$$\mu \geqslant 0$$

and negative semidefinite if

$$\mu \leqslant 0$$

A (symmetric) matrix, B, is said to be positive definite if the relation

$$x^T Bx > 0$$

is satisfied. Similar definitions hold for negative definite, and positive and negative semidefinite matrices.

Example (a).

$$\mu (x_1, x_2) = x_1^2 + x_2^2$$

This is positive definite for all x.

Example (b)

$$\mu\ (x_1, x_2) = x_1^2 + x_2^2 - x_1^3$$

This is positive definite for all x with $\|x\| < 1$.

Example (c)

$$\mu\ (x_1, x_2) = (x_1 - x_2)^2$$

Since $\mu = 0$ whenever $x_1 = x_2$, this is positive semidefinite.

Example (d)

$$\mu\ (x_1, x_2, x_3) = x_1^2 + x_2^2$$

Here $\mu = 0$ whenever $x_1 = x_2 = 0$ and x_3 is finite. Therefore μ is semidefinite.

<u>Definition 11</u>: If the system described by

$$\dot{x} = f(x,t) \tag{80}$$

has the property that

$$f\ (a,\ t) = 0 \quad , \quad t \geqslant 0$$

then a is called an equilibrium point of the system. Whenever an <u>equilibrium point</u> exists, a simple transformation of variables, $z = x - a$, transfers this point to the origin. Consequently there is no loss of generality in assuming that the equilibrium point is always located at the origin.

<u>Definition 12</u>: We consider the autonomous system

$$\dot{x} = f\ (x) \tag{81}$$

which has an equilibrium point at the origin; viz:

$$x(o) = 0 \tag{82}$$

Let $R^{(1)}$ be a region in the state space for which $\|x\| < \alpha$, and let $R^{(2)}$ be a similar region for which $\|x\| < \beta$. We may envision these as hyperspheres centered at the origin with radii α and β respectively. It is assumed that $\alpha > \beta$. Further, let the initial state of the system at time $t = 0$ be $x = b \neq 0$. We say that the system is:

1. Stable, if for every b contained in $R^{(2)}$ there is a region $R^{(3)}$ defined by $\|x\| < \gamma$, where $\alpha > \gamma > \beta$, such that $\lim_{t \to \infty} [x(t)] < \alpha$. In other words, the trajectory of the system never leaves the region $R^{(3)}$.
2. Asymptotically stable, if for any b in $R^{(2)}$, $\lim_{t \to \infty} [x(t)] \to 0$.
3. Asymptotically stable in the large (completely stable), if it is asymptotically stable and the region $R^{(2)}$ encompasses the entire state space.
4. Unstable, if for some b contained in region $R^{(2)}$ with β arbitrarily small, the $\lim_{t \to \infty} [x(t)] > \alpha$.

It will be noted that the concept of stability is not as sharply delineated as in the linear case. The type of stability is generally a function of the initial state. As a rule, a stable system in an engineering sense, is either characterized by a limit cycle of "small" amplitude or else a convergent trajectory to the equilibrium point. When a system is asymptotically stable, it is desirable to know the region of asymptotic stability. Furthermore, it is important to determine the conditions which ensure complete stability. The basic theorems which provide an answer to some of these questions are contained in the following section.

Remark: There is some compromise with mathematical rigor in the above definitions. A more rigorous formulation could not be attempted without requiring a firmer theoretical foundation. The definitions are however eminently satisfactory for engineering purposes -- and this is the theme of the present exposition. Sharper definitions as well as a more extended discussion of various types of stability, together with various existence theorems, may be found in the literature. (18)

3.3.2 Basic Theorems

The very active research, in recent years, in the area of nonlinear stability has produced a rich variety of results for determining the stability of nonlinear systems. Much of this work is of a highly theoretical nature with no immediate application to practical systems. In keeping with the aim of the present exposition, we will present only those results which have proved useful in the study of modern nonlinear control systems. The discussions will be, where ever possible, in the engineering vernacular, and most of the formal presentations will be supplemented with practical (albeit simplified) examples.

We consider an n^{th} order autonomous nonlinear system described by

$$\dot{x} = f(x) \tag{83}$$

with an equilibrium point given by

$$x(o) = 0 \tag{84}$$

In what follows, we shall be concerned with a real valued scalar function (the Lyapunov function) having one or more of the following properties:

I. All the first partial derivatives of $V(x)$ with respect to the components, x_i, of x exist and are continuous for all x.

II. $V(x)$ is positive definite; i.e.,

$$V(x) > 0 \qquad x \neq 0$$

$$V(o) = 0$$

III. The quantity $\dot{V}(x) = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} \frac{dx_i}{dt} = \nabla V \cdot x$

is negative definite; i.e.,

$$\dot{V}(x) < 0 \qquad x \neq 0$$

$$\dot{V}(o) = 0$$

IV. The quantity $\dot{V}(x)$ is negative semidefinite; i.e.,

$$\dot{V}(x) \lesseqgtr 0 \qquad x \neq 0$$

V. $V(x) \rightarrow \infty$ as $||x|| \rightarrow \infty$

VI. $\dot{V}(x) > 0$

The following theorems refer to the system described in Eqs. (83), (84) and the properties of an appropriate Lyapunov function.

Theorem 1: If there exists a Lyapunov function, $V(x)$, satisfying I, II, and IV, then the system (83) is stable in the vicinity of the origin.

Theorem 2: If there exists a Lyapunov function, $V(x)$ satisfying I, II, and III, then the origin in (83) is asymptotically stable.

Theorem 3: If $V(x)$ satisfies I, II, III, and V, then the origin is asymptotically stable in the large.

Theorem 4: If $V(x)$ satisfies I, II, IV, and V, and if in addition $\dot{V}(x)$ is not identically zero along any solution of (83) other than the origin, then the origin is asymptotically stable in the large.

Theorem 5: If $V(x)$ satisfies I, II, and VI, then the origin of (83) is unstable.

Remark: Quite often it is difficult to assure or confirm asymptotic stability in the large. One can establish only that the origin is stable, or else asymptotically stable. These are only local concepts. The previous theorems (1 and 2) give no information on the extent of stability or asymptotic stability. We know only that if the perturbations are not "too large," the system tends to return to the equilibrium state, but we know nothing about what "too large" means. The following theorem due to La Salle[24] helps clarify the situation.

Theorem 6: Let $R^{(r)}$ denote the region where $V(x) < r$. Within this region, assume that $V(x)$ satisfies I and II and that the region $R^{(r)}$ is bounded. Then any solution of Eq. (83) which starts in $R^{(r)}$ is:

a. Stable and remains in $R^{(r)}$ if $V(x)$ satisfies IV in this region.

b. Asymptotically stable if $V(x)$ satisfies III in $R^{(r)}$.

The theorems presented in this section represent only a small fraction of the available theorems relating to Lyapunov's second method. They do include, however, most of those of direct interest to the control engineer. Various additional results and generalizations, applicable in special circumstances, will be presented in the following sections.

A few general observations are in order at this point. We note first of all that the conditions for stability given thus far are sufficient; they may not be necessary. Consequently, failure to satisfy the conditions of these theorems does not mean that the system being analyzed is necessarily unstable. Furthermore, a Lyapunov function for a given system is not, in general, unique. The determination of that Lyapunov function which yields the least restrictive conditions on stability is often a difficult problem.

Finally, the most obvious limitation in the practical application of the direct method is the lack of a general systematic procedure for generating Lyapunov functions. A variety of specialized techniques, applicable in particular circumstances, will be discussed in the sections to follow. The use of the theorems presented thus far will be illustrated in the following examples.

Example 8:

$$\dot{x}_1 = x_2 - \alpha x_1 (x_1^2 + x_2^2)$$

$$\dot{x}_2 = -x_1 - \alpha x_2 (x_1^2 + x_2^2)$$

where $\alpha \equiv$ positive constant.

If we take

$$V = x_1^2 + x_2^2$$

then

$$\dot{V} = -2\alpha (x_1^2 + x_2^2)^2$$

which is obviously negative definite for all x. Furthermore, we note that

$$V \to \infty \quad \text{as } \|x\| \to \infty$$

Consequently, by theorem 3, the system is asymptotically stable in the large.

Example 9:

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -\alpha (1 + x_2)^2 x_2 - x_1$$

$$\alpha \equiv \text{positive constant}$$

Again take

$$V = x_1^2 + x_2^2$$

It follows that

$$\dot{V} = -2\alpha (1 + x_2)^2 x_2^2$$

which is negative semidefinite*. Also

*Since $\dot{V} = 0$ for $x_2 = 0$ and x_1 arbitrary.

$V \to \infty \quad \text{as } \|x\| \to \infty$

We will now show that $\dot{V} = 0$ is not a trajectory of the system. In fact, $\dot{V} = 0$ implies that x_2 equals zero or -1. The slope of the system trajectory is

$$\frac{dx_2}{dx_1} = -\alpha(1 + x_2)^2 - \frac{x_1}{x_2}$$

But for $x_2 = 0$, this slope is infinite, which means that the x_1 axis cannot be a trajectory of the system. Furthermore, for $x_2 = -1$, we have

$$\frac{dx_2}{dx_1} = x_1$$

which is never zero except at the origin. Therefore the line $x_2 = -1$ cannot be a system trajectory.

Since we have now fulfilled the conditions of theorem 4, we conclude that the origin of the given system is asymptotically stable in the large.

<u>Example 10</u>:

$$\ddot{\epsilon} + (1 - |\epsilon|)\dot{\epsilon} + \epsilon = 0$$

Putting $\epsilon = x_1$, we write this as

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -(1 - |x_1|)x_2 - x_1$$

Again taking

$$V = x_1^2 + x_2^2$$

we find

$$\dot{V} = -2x_2^2(1 - |x_1|)$$

Consider now the region in the state space which is bounded by the curve

$$x_1^2 + x_2^2 = 1$$

Within this region V is positive definite and $\dot{V}$ is negative definite. Therefore, by theorem 6, this region is asymptotically stable.

Example 11:

$$\ddot{x}_1 + \dot{x}_1 + x_1^3 = 0$$

We write this as

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -x_2 - x_1^3$$

Here, if we take $V = x_1^2 + x_2^2$ as before, we can obtain no conclusive results. However, if we use

$$V = \frac{1}{4} x_1^4 + \frac{1}{2} x_2^2$$

then there follows

$$\dot{V} = - x_2^2$$

which is negative semidefinite. Reasoning, analogous to that of Example 2, shows that $\dot{V} = 0$ cannot be a trajectory of the system. And since $V \to \infty$ as $\|x\| \to \infty$, theorem 4 enables us to conclude that the origin of the system is asymptotically stable in the large.

Example 12: Some of the examples considered thus far can be analyzed from a more unified point of view.[19] Consider the dynamical system

$$\ddot{\epsilon} + g(\epsilon)\dot{\epsilon} + h(\epsilon) = 0$$

We assume that g and h are polynomials in ϵ, with g even and h odd. It is also assumed that $h(o) = 0$. Define

$$\varphi(\epsilon) = \int_0^{\epsilon} g(\epsilon)\, d\epsilon$$

$$\psi(\epsilon) = \int_0^{\epsilon} h(\epsilon)\, d\epsilon$$

Note that φ is odd and ψ even, and that

$$\varphi(o) = \psi(o) = 0$$

Now write $x_1 = \epsilon$, and express the system in the equivalent form

$$\dot{x}_1 = x_2 - \varphi(x_1)$$

$$\dot{x}_2 = -h(x_1)$$

Note that under the foregoing assumptions, the origin is an equilibrium point of the system.

As a Lyapunov function, we take

$$V = \frac{1}{2} x_2^2 + \psi(x_1)$$

It follows that

$$\dot{V} = -h(x_1)\, \varphi(x_1)$$

We now suppose that there exist two positive constants, α and β, such that

$$h(x_1)\, \varphi(x_1) > 0 \quad \text{for } |x_1| < \alpha, \quad x_1 \neq 0$$

$$\psi(x_1) < \beta \text{ implies } |x_1| < \alpha$$

If these two relations are satisfied, then the region defined by $V < \beta$ is bounded and within this region $\dot{V} \leqq 0$. Furthermore $\dot{V} = 0$ in this region only when $x_1 = 0$. But the x_2 axis is not a trajectory of the system since the slope

$$\frac{dx_2}{dx_1} = -\frac{h(x_1)}{x_2 - \varphi(x_1)}$$

is finite for $x_2 = 0$. By theorem 6, we conclude that every solution initiating in the interior of the region defined above is asymptotically stable.

As an example of this approach consider the system

$$\ddot{\epsilon} + \mu (1 - \epsilon^2) \dot{\epsilon} + \epsilon = 0$$

where $\mu \equiv$ positive constant. In this case, with $\epsilon = x_1$, we have

$$\dot{x}_1 = x_2 - \varphi (x_1)$$

$$\dot{x}_2 = -h (x_1)$$

with

$$g (x_1) = \mu (1 - x_1^2)$$

$$h (x_1) = x_1$$

$$\varphi (x_1) = \mu \left(x_1 - \frac{x_1^3}{3} \right)$$

$$\psi (x_1) = \frac{1}{2} x_1^2$$

Taking

$$V = \frac{1}{2} x_2^2 + \psi (x_1)$$

$$= \frac{1}{2} (x_2^2 + x_1^2)$$

we find

$$\dot{V} = - \mu x_1^2 \left(1 - \frac{x_1^2}{3} \right)$$

which is negative semidefinite for $|x_1| < \sqrt{3}$. Consequently, we take $\alpha = \sqrt{3}$. Now if $\psi (x_1) < \beta$ is to imply that $|x_1| < \sqrt{3}$, we must have $\beta = 3/2$. It follows, therefore, that the region bounded by $x_1^2 + x_2^2 = 3$ is a region of asymptotic stability.

3.3.3 Lur'e Method

The earliest successful attempt to evolve a systematic method for generating Lyapunov functions is due to Lur'e.[20] This technique is the subject of a book by Letov[21] which discusses many applications in great detail, but which is overly burdened by an awkward notation. As a matter of fact, it is possible to exhibit all the essential features of the method quite clearly and succinctly by adopting matrix notation. Various distinctions considered by Letov and other Russian investigators are shown to be irrelevant. The emphasis in the present discussion will be on adapting the basic results to realistic control systems without, however, compromising the analytical development. The power of the Letov method lies in its ability to analyze systems of moderate order. There are, nevertheless, serious restrictions on the type of nonlinear system which may be treated. To motivate the discussion, we consider the system shown in Fig. 51. The nonlinear function, $\varphi(\sigma)$, satisfies the conditions

$$\int_0^{\sigma} \varphi(\sigma)\, d\sigma \gtrless 0 \qquad \text{for all } |\sigma| > 0$$
$$\varphi(o) = 0 \tag{85}$$

G(s) is a linear transfer function of the form

$$G(s) = \frac{\theta}{\varphi} = \frac{\sum_{i=0}^{n} \beta_{n-i}\, s^i}{\sum_{i=0}^{n} \alpha_{n-i}\, s^i} \tag{86}$$

We assume that the denominator of this expression contains no multiple or zero roots and that each of the roots has a negative real part. It may also be assumed, without loss of generality, that the coefficient of s^n in the denominator is unity. We also add the restriction that the degree of the numerator is at least one less than that of the denominator. In summary, it is stipulated that

$$\alpha_o = 1 \qquad \beta_o = 0 \tag{87}$$

Under these conditions, it is always possible to write the input-output dynamics of the linear element in the form[25]

$$\dot{x} = Ax + b\,\varphi(\sigma)$$
$$\sigma = -x_1 \tag{88}$$

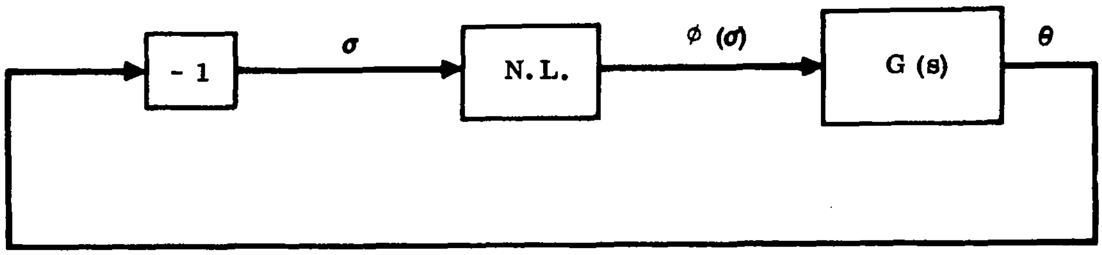


Figure 51. Representation of Nonlinear System

where x and b are n vectors and A is a constant matrix determined from

$$\theta = x_i$$

$$\dot{x}_i = x_{i+1} + b_i \varphi \tag{89}$$

$$i = 1, 2, \ldots\ldots\ldots\ldots, i-1$$

$$\dot{x}_n = -\sum_{i=1}^{n} \alpha_{n-i+1} x_i + b_n \varphi \tag{90}$$

$$b_o = \beta_o = 0$$

$$b_i = \beta_i - \sum_{j=0}^{i-1} \alpha_{i-j} b_j \tag{91}$$

$$i = i, 2, \ldots\ldots\ldots\ldots, n$$

This method is illustrated by a simple example as follows.

$$G(s) = \frac{\theta}{\varphi} = \frac{\beta_1 s + \beta_2}{s^2 + \alpha_1 s + \alpha_2}$$

From Eq. (91) we find

$$b_1 = \beta_1 - \alpha_1 b_o = \beta_1$$

$$b_2 = \beta_2 - \alpha_2 b_o - \alpha_1 b_1 = \beta_2 - \alpha_1 \beta_1$$

Also, using Eqs. (89) and (90)

$$\dot{x}_1 = x_2 + \beta_1 \varphi$$

$$\dot{x}_2 = - \alpha_2 x_1 - \alpha_1 x_2 + (\beta_2 - \alpha_1 \beta_1) \varphi$$

Expressed in matrix notation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\alpha_2 & -\alpha_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \varphi$$

$$\theta = x_1$$

Notice that in the format of Eq. (88), the numerator dynamics (derivatives of the driving function) are effectively eliminated. In the analysis to follow, we will consider a slightly generalized version of Eq. (88); namely,

$$\begin{aligned} \dot{x} &= Ax + b\varphi(\sigma) \\ \sigma &= v^T x \end{aligned} \tag{92}$$

(v $\equiv$ a constant n vector)

Lur'e considered the dynamical system

$$\begin{aligned} \dot{x} &= Ax + b\varphi(\sigma) \\ \dot{\xi} &= \varphi(\sigma) \\ \sigma &= v^T x + \gamma\xi \end{aligned} \tag{93}$$

He distinguished between <u>indirect control</u> (when $\gamma \neq 0$) and <u>direct control</u> ($\gamma = 0$). We will in fact show that these distinctions are irrelevant and that Eqs. (93) are completely equivalent to Eqs. (92).

Suppose that $\gamma \neq 0$. Differentiating the third of Eqs. (93) and making use of the first two, we find

$$\begin{aligned} \dot{\sigma} &= v^T \dot{x} + \gamma \dot{\xi} \\ &= v^T Ax + (v^T b + \gamma)\varphi \end{aligned}$$

There is indeed no loss of generality in assuming that $\gamma = 0$. For if $\gamma \neq 0$, and we write, instead of Eqs. (93)

$$\begin{aligned} \dot{x} &= Ax + b\varphi + f\varphi \\ \sigma &= v^T x \end{aligned}$$

where f is a constant vector defined by

$$\gamma = v^T f$$

one finds that

$$\begin{aligned} \dot{\sigma} &= v^T \dot{x} \\ &= v^T (Ax + b\varphi + f\varphi) \\ &= v^T Ax + (v^T b + \gamma)\varphi \end{aligned}$$

Consequently a nonzero γ need not be considered explicitly if the b vector is appropriately modified.

All of the analysis to follow will deal with the system of Eqs. (92), with the nonlinearity satisfying conditions (85). By virtue of the stipulations placed on the linear element, G(s), the constant matrix, A, satisfies the following.

I. All the eigenvalues of A have negative real parts.

II. No eigenvalue of A equals zero.

III. All the eigenvalues of A are distinct.

Some of these restrictions will be relaxed in Section 3.3.4.

The system (92), which is in the phase variable (canonical) form, may be expressed in diagonal form by making the transformation

$$x = Ty \tag{94}$$

where

$$T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix} \tag{95}$$

This matrix is known as the Vandermonde matrix; the λ_i are the eigenvalues of A. As a result of this operation, we find

$$\begin{aligned} \dot{y} &= Fy + m\varphi(\sigma) \\ \sigma &= c^T y \end{aligned} \tag{96}$$

where

$$F = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \equiv T^{-1} A T \tag{97}$$

$$m = T^{-1} b \tag{98}$$

$$c = T^{T} v \tag{99}$$

For the Lyapunov function, we take*

$$V = y^{T} P y + \int_{0}^{\sigma} \varphi(\sigma)\, d\sigma \tag{100}$$

Taking the derivative with respect to time using Leibnitz's formula, and making use of Eq. (96), with $\dot{\sigma} = c^T \dot{y}$ one obtains

$$\dot{V} = \dot{y}^{T} P y + y^{T} P \dot{y} + \varphi(\sigma) \frac{d\sigma}{dt} + \int_{0}^{\sigma} \frac{\partial \varphi(\sigma)}{\partial t}\, d\sigma$$

$$= \dot{y}^{T} P y + y^{T} P \dot{y} + \dot{\sigma} \varphi(\sigma)$$

$$= - y^{T} Q y + \varphi\, (2m^{T} P + c^{T} F)\, y + \varphi^{2} c^{T} m \tag{101}$$

where

$$Q = - (F^{T} P + PF) \tag{102}$$

In seeking to establish asymptotic stability in the large, we must have (by theorem 3)

$$V > 0$$

$$\dot{V} < 0$$

$$V \longrightarrow \infty \quad \text{as} \quad \|y\| \longrightarrow \infty$$

If P is positive definite, and if (85) holds, the first and third conditions are automatically satisfied. To satisfy the second condition, Q must be positive definite. It is known[19] that if all the eigenvalues of F have negative real parts, then in Eq. (102) if P is positive definite so is Q, and vice versa. One may therefore choose <u>either</u> P or Q in some appropriate manner.

Following Lur'e, we let

$$Q = aa^{T} \tag{103}$$

The components of the vector, a, are chosen as follows. If $\lambda_i = \lambda_j^*$, where ( )* is used to denote complex conjugate, then $a_i = a_j^*$. Otherwise, the a_i's are

*P is a constant symmetric matrix whose elements are, for the moment, undetermined.

simply constants (as yet undetermined). Writing

$$r = -c^T m \tag{104}$$

and substituting Eq. (103) into Eq. (101), and then completing the square, results in

$$\dot{V} = -\left[a^T y - \varphi\sqrt{r}\right]^2 + \varphi\left[2m^T P + c^T F + 2\, a^T\sqrt{r}\right] y \tag{105}$$

If $r > 0$ and $\varphi(o) = 0$, then the above expression is negative definite if

$$2m^T P + c^T F + 2a^T\sqrt{r} = 0 \tag{106}$$

Consequently, if a vector, a, can be found to satisfy this equation, we may assert that the system (92) is asymptotically stable in the large (via theorem 3).

To solve Eq. (106), it is necessary to solve Eq. (102) for P, where Q is given by Eq. (103). This solution is[27]

$$p = \int_o^\infty e^{F^T t} Q\, e^{Ft}\, dt \tag{107}$$

Now

$$e^{Ft} = \begin{bmatrix} e^{\lambda_1 t} & & 0 \\ & e^{\lambda_2 t} \ddots & \\ 0 & & e^{\lambda_n t} \end{bmatrix}$$

It follows therefore that the $(ij)^{th}$ component of the matrix

$$e^{F^T t}\; Q\; e^{Ft}$$

is

$$a_i\, a_j\, e^{(\lambda_i + \lambda_j)t}$$

Furthermore

$$a_i a_j \int_0^\infty e^{(\lambda_i+\lambda_j)t}\, dt = -\frac{a_i a_j}{(\lambda_i + \lambda_j)}$$

Therefore, the $(ij)^{th}$ component of the P matrix is given by

$$P_{ij} = -\frac{a_i a_j}{(\lambda_i + \lambda_j)} \tag{108}$$

Various additional stability criteria may be developed by taking modified forms of the Lyapunov function, (100), and placing additional restrictions on the nonlinearity. The reader is referred to Letov[21] for details. We now illustrate the application of this method by a simple example.

Example 13: The system block diagram is as shown in Fig. 51, and the nonlinearity satisfies (85). The linear transfer function is

$$G(s) = \frac{(s+4)}{(s+1)\ (s+2)} = \frac{\theta}{\varphi}$$

$$= \frac{s + 4}{s^2 + 3s + 2}$$

Using relations (89) through (91), we obtain

$$\dot{x} = Ax + b\varphi$$

where

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$\sigma = v^T x = -x_1 = -\theta_1$$

where

$$v = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

The eignevalues of A are, in general, obtained from ($I \equiv$ identity matrix)

$$|A - \lambda I| = 0$$

However, in the present case, these eigenvalues are the poles of the open-loop transfer function, G(s); viz.

$$\lambda_1 = -1$$

$$\lambda_2 = -2$$

Calculating the Vandermonde matrix

$$T = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$$

and its inverse

$$T^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

we then have

$$F = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}$$

from Eq. (97), while

$$m = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

$$c = -\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

from Eqs. (98) and (99) respectively. Furthermore, via Eq. (104), we find

$$r = -c^T m = 1$$

and

$$p = \frac{1}{12} \begin{bmatrix} 6\, a_1^2 & 4\, a_1\, a_2 \\ 4\, a_1 a_2 & 3\, a_2^2 \end{bmatrix}$$

using Eq. (108).

Finally, in order to satisfy Eq. (106), and therefore to establish that the system is completely stable (asymptotically stable in the large), there must exist real constants, a_1 and a_2, satisfying

$$a_1^2 + \frac{1}{3}(2 - \frac{4}{3} a_2)\, a_1 + \frac{1}{3} = 0$$

$$a_1^2 - 2\,(a_1 + 1)\, a_2 - 2 = 0$$

This is indeed the case since we find that $a_1 = -0.98$ and $a_2 = -1.41$.

For higher order systems, it becomes increasingly more awkward to solve Eq. (106) for the vector a. For an n^{th} order system, the n components of the vector, a, are to be obtained by solving n simultaneous algebraic equations. This presents serious difficulties at the computational level. An approach which offers some advantages in this respect is the following.

We assume that

$$Q = \begin{bmatrix} \gamma_1 & & & 0 \\ & \gamma_2 & & \\ & & \ddots & \\ 0 & & & \gamma_n \end{bmatrix} \tag{109}$$

Here the γ_i's are positive constants. Substituting this expression for Q in Eq. (101), and completing the square, yields

$$-\dot{V} = (y + \varphi Q^{-1} u)^T Q (y + \varphi Q^{-1} u) + (r - u^T Q^{-1} u) \varphi^2 \tag{110}$$

where

$$u = -(Pm + \frac{1}{2} Fc) \tag{111}$$

and r is as defined by Eq. (104). Since Q as defined by Eq. (109) is obviously positive definite, the right-hand side of Eq. (110) is positive if

$$r > u^T Q^{-1} u \tag{112}$$

This condition replaces (106) as the requirement for complete stability. The essential difference is that (112) is a scalar inequality rather than a set of simultaneous algebraic equations.

A slight complication arises when some of the eigenvalues of the state matrix, A, are complex. For γ_i in Eq. (109) positive, this will result in some of the components of P, as calculated from Eq. (102), being complex, with the result that some of the components of u, obtained from Eq. (111), are also complex. This simply means that in (112) one simply uses the Hermitian form, $u^* Q^{-1} u$ rather than $u^T Q^{-1} u$. Nothing else is changed.

The following example illustrates the application of the method.

Example 14: We consider the system whose nonlinearity satisfies (85) and which is described by

$$\dot{x} = Ax + b\varphi$$

$$\sigma = v^T x$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -30 & -31 & -10 \end{bmatrix}$$

$$b = \begin{bmatrix} 0 \\ 1 \\ -4 \end{bmatrix}$$

$$v = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$$

The eigenvalues of A are

$$\lambda_1 = -2$$

$$\lambda_2 = -3$$

$$\lambda_3 = -5$$

so that

$$F = \begin{bmatrix} -2 & & 0 \\ & -3 & \\ 0 & & -5 \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 1 & 1 \\ -2 & -3 & -5 \\ 4 & 9 & 25 \end{bmatrix}$$

$$T^{-1} = \frac{1}{6} \begin{bmatrix} 30 & 16 & 2 \\ -30 & -21 & -3 \\ 6 & 5 & 1 \end{bmatrix}$$

and

$$m = T^{-1}b = \frac{1}{6}\begin{bmatrix} 8 \\ -9 \\ 1 \end{bmatrix}$$

$$c = T^T v = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

$$r = -c^T m = 8$$

Taking Q as

$$Q = \begin{bmatrix} \gamma_1 & & 0 \\ & \gamma_2 & \\ 0 & & \gamma_3 \end{bmatrix}$$

where the γ_i's are (as yet undetermined) positive constants, we solve for P either from Eq. (107) or (102). The result is

$$P = \begin{bmatrix} -\frac{1}{2}\frac{\gamma_1}{\lambda_1} & & 0 \\ & -\frac{1}{2}\frac{\gamma_2}{\lambda_2} & \\ 0 & & -\frac{1}{2}\frac{\gamma_3}{\lambda_3} \end{bmatrix}$$

The components of the vector, u, are now obtained from Eq. (111); viz.

$$u_i = \frac{1}{2\lambda_i}(m_i \gamma_i + \lambda_i)^2$$

$$i = 1, 2, 3$$

where m_i denotes the i^{th} component of the vector, m.

Consequently, the inequality (112) becomes

$$8 \geqslant \frac{(\gamma_1 + 3)^2}{9\gamma_1} + \frac{(\gamma_2 - 6)^2}{16\gamma_2} + \frac{(\gamma_3 + 100)^2}{100\gamma_3}$$

One set of positive γ_i's which satisfies this inequality is

$$\gamma_1 = 1$$

$$\gamma_2 = 6$$

$$\gamma_3 = 50$$

We conclude therefore that the given system is completely stable.

The Lur'e method represents an important step in the development of a fairly general theory of stability for nonlinear systems.

The technique can be applied to moderately high order systems described by the format of Eqs. (92) with the nonlinearity satisfying (85). A more detailed description of the nonlinear element is not required. On the other hand, phase plane methods are limited to second order systems, while the describing function technique requires a fairly precise description of the nonlinearity.

There are several severe limitations, however. It is possible that the criteria of Eqs. (106) or (112) will fail to be satisfied even if the system is stable. In this case various modifications of the assumed Lyapunov function, (100), may be taken, together with the application of additional restrictions on the nonlinearity. This will yield slightly different stability criteria which may be satisfied where others might not. Several such approaches are considered by Letov.[21]

Nevertheless, there are restrictions of a fundamental nature such that stable systems will always be rejected by the Luré method. Consider, for example, a linear element described by

$$\varphi = k\sigma$$

$$(0 < k < \infty)$$

This obviously satisfies (85). It is apparent therefore, that in the case of a linearized system, the Lur'e stability criteria would select as stable only those systems that are stable for all positive values of the open-loop gain. If the root locus of the open-loop transfer function, G(s), is not confined to the left half of the s plane, a linearized system will, for some positive value of the open-loop gain, be unstable. Hence the stability criteria developed in this section will reject all those systems of the type shown in Fig. 52.

It must be emphasized, however, that the fact that a system with a single nonlinear gain element is confined to the left half of the s plane does not imply that the system is stable. Conversely, the fact that the linear portion of a system with a single nonlinear gain element is not confined to the left half of the s plane does not imply that the system is unstable.

Based on the foregoing considerations and the assumptions inherent in the Lur'e method, we may summarize the reasons why stable systems will be rejected by this method as follows.

a. The root locus of the linear part of the system is not confined to the left half of the s plane.

b. There are open-loop poles at the origin of the s plane.

c. The open-loop function has multiple poles.

d. The constant, r, in (106) or (112) is negative.

An approach which broadens the scope of the Lur'e method considerably, and which overcomes most of the above limitations is described in the following section.

3.3.4 Pole and Zero Shifting

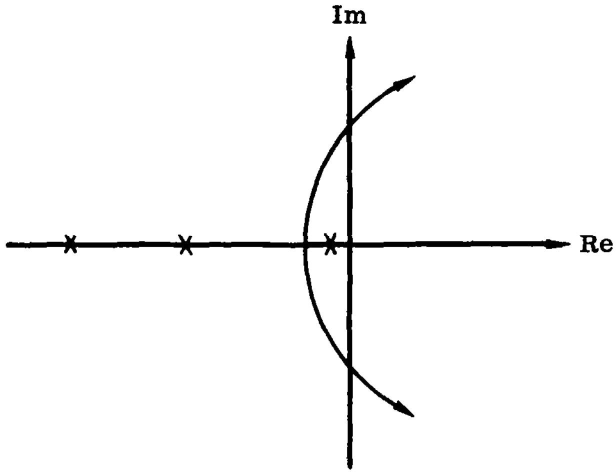
Some of the most serious limitations in the Lur'e method may be removed by adopting an approach developed by Rekasius and Gibson.[22] The pole shifting technique may be used to:

a. Analyze a system whose open-loop gain does not fall below a certain minimum value.

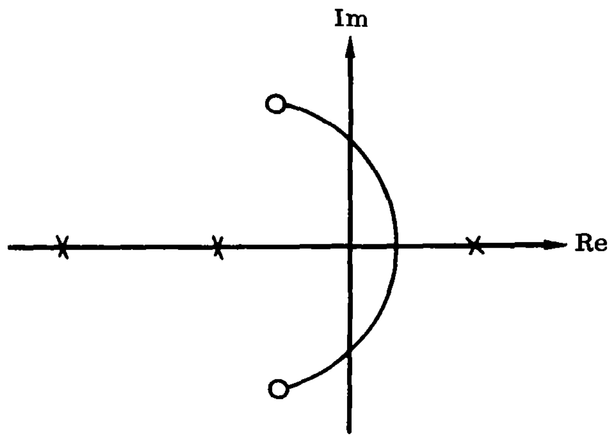
b. Separate multiple open-loop poles.

c. Analyze systems with open-loop poles in the right half plane.

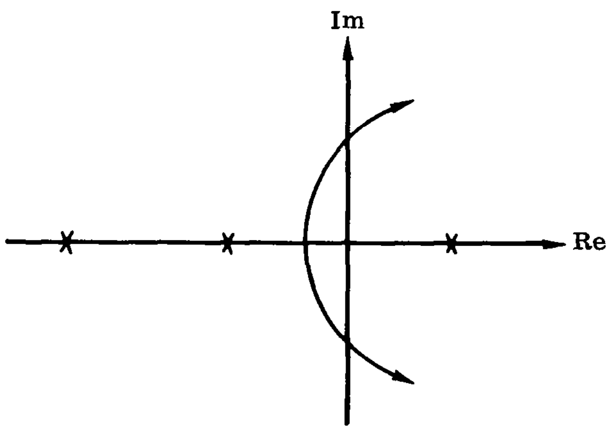
Essentially, what is involved is an appropriate change of variable such that the transformed system satisfies the restrictions of the Lur'e method. We will



a. Stable for low values of gain



b. Stable for high values of gain



c. Stable for intermediate values of gain

Figure 52. Root Loci for Certain Linear Systems

consider the system depicted in Fig. 51 and described by Eq. (92). It is assumed that the nonlinearity satisfies the inequality

$$\sigma\varphi > \alpha\,\sigma^2 \tag{113}$$

where α is a positive constant.

In other words, the nonlinearity is confined to the first and third quadrants such that $|\varphi|$ always exceeds $\alpha\,|\sigma|$. (See Fig. 53.) Now define a new variable, φ', by

$$\varphi' = \varphi - \alpha\sigma \tag{114}$$

Substituting this in (113) yields the new condition

$$\sigma\varphi' > 0 \tag{115}$$

which obviously satisfies (85). The revised block diagram is shown in Fig. 54. If we now deal with the new nonlinear parameter, φ', defined by Eq. (114) and the new linear transfer function

$$\bar{G}(s) = \frac{G(s)}{1 + \alpha\,G(s)} \tag{116}$$

it is found that all of the stipulations of the Lur'e method may be satisfied. Specifically, we note that if

$$G(s) = \frac{N(s)}{D(s)}$$

where N(s) and D(s) are polynomials in s, then

$$\bar{G}(s) = \frac{N(s)}{D(s) + \alpha N(s)}$$

It is obvious that an arbitrarily small value of α will separate multiple poles and that increasing values of α, will, in most cases of interest, shift a pole in the right half plane into the left half plane. Consequently the pole shifting technique may be used to prove the stability of systems where the Lur'e method fails.

The zero shifting technique may be employed to study the stability of systems whose open-loop gain does not exceed some specified maximum. In particular, suppose that the nonlinearity satisfies the inequality

$$0 \leqslant \sigma\varphi < \beta\sigma^2 \tag{117}$$

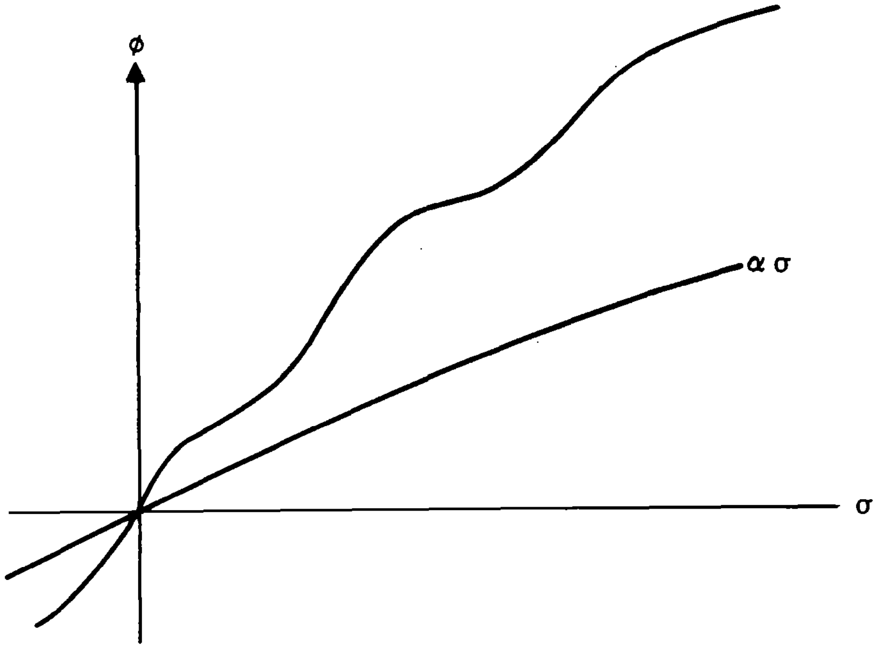


Figure 53. Characteristics of Nonlinear Element

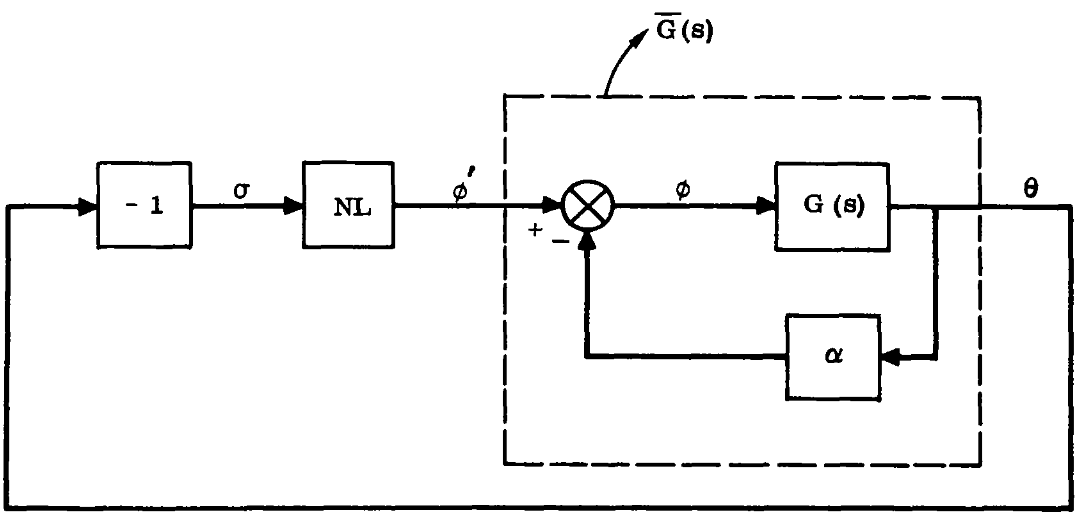


Figure 54. Revised Block Diagram in Pole Shifting Method

where β is a positive constant. (See Fig. 55.)

Defining a new variable by

$$\sigma' = \sigma \frac{\varphi}{\beta} \tag{118}$$

and substituting back in (117) yields

$$\sigma' \varphi \geqslant 0$$

However, the Lur'e method cannot be applied directly since the input to the non-linear element, σ', is no longer given by the second of Eqs. (92). (See Fig. 56.) A Lyapunov function other than (100) must be tried. If we select

$$V = y^T Py$$

where y represents the canonic state vector in the system described by (96), and P is a constant symmetric matrix, then we find

$$\dot{V} = - y^T Qy + 2\varphi m^T Py$$

where Q is given by (102). As in the Lur'e method, we take Q to be of the form (103). Completing the square in $\dot{V}$ then yields

$$\dot{V} = - (a^T y)^2 + \varphi (2m^T P) y$$

The resulting stability equation becomes

$$m^T P = 0 \tag{119}$$

By using pole shifting combined with zero shifting, an even wider range of systems may be analyzed. Reference 22 contains a variety of examples together with several modified Lyapunov functions.

3.3.5 The Variable Gradient Method

The usual methods of Lyapunov stability analysis begin with an assumption of a tentative Lyapunov function from which an attempt is made to determine those conditions which ensure that the system is, in some sense, stable. The selection of a suitable Lyapunov function, appropriate to a given system, is a task which generally taxes the ingenuity of the analyst. In this section, we consider methods for generating Lyapunov functions systematically. Various studies in this connection have been reported by Szego(28) and Ingwerson(29), but the variable gradient method, developed by Schultz and Gibson, (23) is by far the most powerful. We will consider this method in some detail.

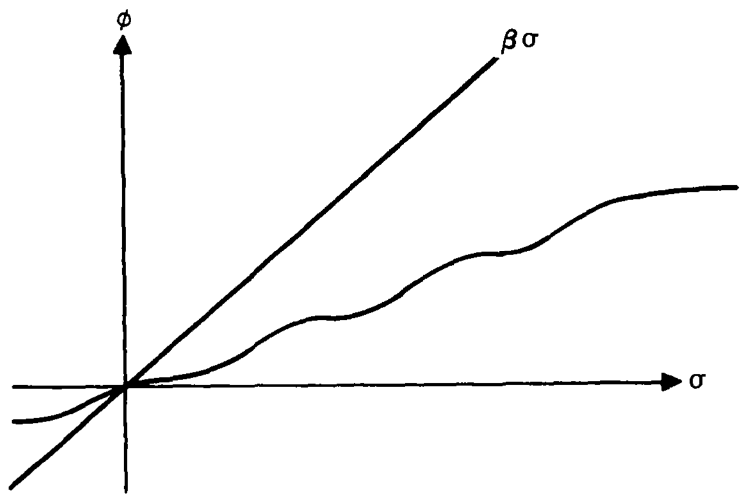


Figure 55. Characteristics of Nonlinear Element

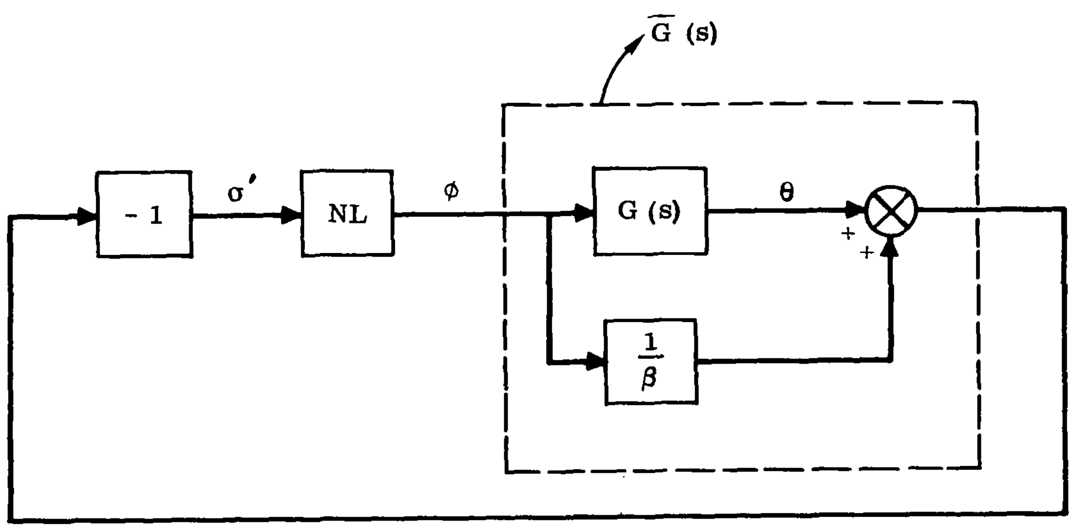


Figure 56. Revised Block Diagram in Zero Shifting Method

It is known that if a system is stable in state space, then a Lyapunov function, V, exists for this system.[18] Furthermore, if V exists, then its gradient, ∇V, also exists. Given this gradient, both V and $\dot{V}$ may be calculated. We note that when V is given one may write

$$\dot{V} = \sum_i \frac{\partial V}{\partial x_i} \frac{dx_i}{dt}$$

where for $\dot{x}_i$, we insert the system equations. Now this may be expressed as

$$\dot{V} = (\nabla V)^T \dot{x} \tag{120}$$

To find V from ∇V, we write

$$V = \int_0^x (\nabla V)^T dx = \int_0^x \nabla V \cdot dx \tag{121}$$

This is a line integral to an arbitrary point in the state space and is independent of the path of integration. The simplest such path is the following

$$V = \int_0^{x_1} V_1 (\gamma_1, 0, \ldots\ldots, 0)\, d\gamma_1 + \int_0^{x_2} V_2 (x_1, \gamma_2, 0, \ldots\ldots, 0)\, d\gamma_2$$

$$+ \ldots\ldots\ldots + \int_0^{x_n} V_n (x_1, x_2, \ldots\ldots, x_{n-1}, \gamma_n)\, d\gamma_n \tag{122}$$

In this expression, V_i ( ) stands for the i^{th} component of the vector, ∇V.

To obtain a unique scalar function, V, by a line integration of a vector function, ∇V, it is necessary that[30]

$$\nabla \times \nabla V = 0 \tag{123}$$

The core of the variable gradient method is contained in the following assumption for the form of the vector, ∇V.

$$
\nabla V = \begin{bmatrix} \alpha_{11} x_1 + \alpha_{12} x_2 + \cdots\cdots + \alpha_{1n} x_n \\ \alpha_{21} x_1 + \alpha_{22} x_2 + \cdots\cdots + \alpha_{2n} x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{n1} x_1 + \alpha_{n2} x_2 + \cdots\cdots + \alpha_{nn} x_n \end{bmatrix} \tag{124}
$$

The coefficients, α_{ij}, are assumed to be made up of a constant portion and a portion which is a function of the state variables; viz.

$$
\alpha_{ij} = \alpha_{ijk} + \alpha_{ij\mu}(x) \tag{125}
$$

Several significant properties of the α_{ij} emerge upon examination of (124). First of all, the solution of a given problem may require that the i^{th} component of ∇V; i.e.

$$
V_i = \alpha_{i1} x_1 + \alpha_{i2} x_2 + \cdots\cdots + \alpha_{in} x_n
$$

contain terms that have more than one state variable as factors. Evidently such terms may be determined from terms such as $\alpha_{ij}(x)\, x_i$; therefore, the $\alpha_{ii\mu}(x)$ may be taken without loss of generality as $\alpha_{ii\mu}(x_i)$. Furthermore, for V to be positive definite in the neighborhood of the origin, α_{iik} must always be positive. Also, if the relation

$$
\lim_{\|x\| \to \infty} V(x) \longrightarrow \infty
$$

is to be satisfied, then $\alpha_{ii\mu}(x_i)$ must be an even function of x_i and greater than zero for large x_i. In case $\alpha_{iik} = 0$, then, for the same reason, we must have $\alpha_{ii\mu}(x_i)$ even and greater than zero for all x_i. Additional restrictions on the α_{ij} are imposed by the curl equations (123). Finally, the α_{ij} must be so chosen, as to make $\dot{V}$ at least negative semidefinite.

In practice, the procedure is often simpler to apply than the above exposition would indicate. To begin with, one may assume the α_{ij}'s are constants, using state dependent α_{ij}'s only if needed to satisfy the curl equations or to make $\dot{V}$ negative semidefinite.

The formal application may be summarized as follows:

a. Assume ∇V of the form shown in Eq. (124).

b. From ∇V determine $\dot{V}$ via Eq. (120) and the state equations of the system.

c. Constrain $\dot{V}$ to be at least negative semidefinite.

d. Determine the remaining coefficients, α_{ij} via Eq. (123).

e. Recheck $\dot{V}$ since step d. may have altered the $\dot{V}$ determined in step c.

f. Determine V via Eq. (122).

g. Check system stability by any of the theorems of Section 3. 3. 2

We illustrate this procedure by a simple example.

Example 15: The system to be analyzed is shown in Fig. 57. The nonlinear element is described by

$$\varphi = \sigma \psi(\sigma)$$

where $\psi(\sigma)$ is some arbitrary function of σ We seek to determine the restrictions on $\psi(\sigma)$ and the zero, β, of the linear transfer function, such that the system is asymptotically stable in the large. This particular problem is treated by Schultz and Gibson,[23] and among their derived stability conditions is that the slope of the nonlinearity must be positive. It will be shown here that this requirement is, in fact, unnecessary. The differing results stem from the fact that the equations of Schultz and Gibson contain derivatives in the forcing function, which may be eliminated by making the transformation of variables described in Section 3. 3. 3. Proceeding in this way, we write the state equations as

$$\dot{x}_1 = x_2 - h(x_1)\, x_1$$

$$\dot{x}_2 = -x_2 - (\beta-1)\, h(x_1)\, x_1$$

where

$$\theta = x_1 = -\sigma$$

$$h(x_1) \equiv \psi(-x_1)$$

From Eq. (124), we have

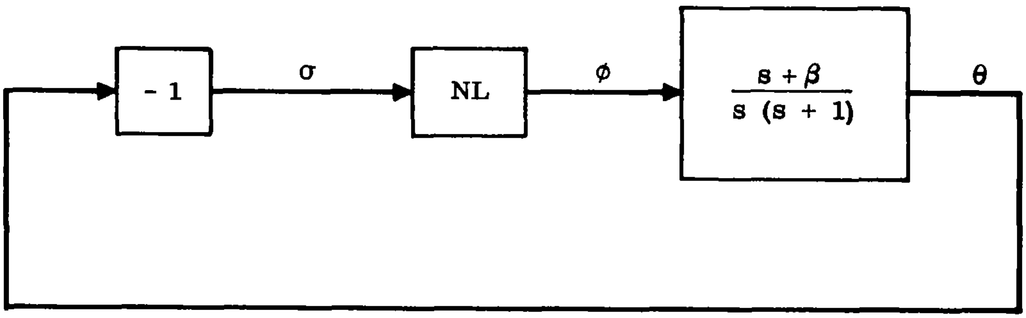


Figure 57. Control Loop for Example 15

$$\nabla V = \begin{bmatrix} \alpha_{11} x_1 + \alpha_{12} x_2 \\ \alpha_{21} x_1 + 2 x_2 \end{bmatrix}$$

where we have arbitrarily set $\alpha_{22} = 2$. Experience has shown that this particular selection for α_{nn} is appropriate for most systems of practical interest. From Eq. (120), we find

$$\dot{V} = \left[\alpha_{11} - \alpha_{21} - \psi(\sigma)\,\alpha_{12} - 2\,\psi(\sigma)\,(\beta - 1)\right] x_1 x_2$$
$$- \psi(\sigma)\left[\alpha_{11} + (\beta - 1)\,\alpha_{21}\right] x_1^2 - (2 - \alpha_{12})\, x_2^2$$

To satisfy (123), we must have

$$\frac{\partial V_1}{\partial x_2} = \frac{\partial V_2}{\partial x_1}$$

or

$$\alpha_{12} = \alpha_{21}$$

If we put

$$\alpha_{12} = 0$$

and

$$\alpha_{11} = 2\,(\beta - 1)\,\psi\,(\sigma)$$

then the expression for $\dot{V}$ becomes

$$\dot{V} = -2\,x_2^2 - 2\,(\beta - 1)\,\psi^2\,(\sigma)\,x_1^2$$

which is obviously negative definite if $\beta > 1$. We now have

$$\nabla V = \begin{bmatrix} 2\,\psi\,(\sigma)\,(\beta - 1)\,x_1 \\ 2\,x_2 \end{bmatrix}$$

Determining V via Eq. (122), it is found that

$$V = x_2^2 + 2\,(\beta - 1)\int_0^{x_1} h\,(x_1)\,x_1\,dx_1$$

Now for $V > 0$ and $\dot{V} < 0$, it is sufficient that

$$\beta > 1$$

$$h\,(x_1)\,x_1 > 0$$

Furthermore, the condition

$$\lim_{\|x\| \to \infty} V\,(x) \to \infty$$

is obviously satisfied. Theorem 3 of Section 3.3.2 enables us to conclude, therefore, that the given system is asymptotically stable in the large if

$$\beta > 1$$

and

$$h(x_1)\, x_1 > 0$$

3.3.6 Miscellaneous Methods

The specific means of applying the Lyapunov theory discussed in the previous sections are perhaps the simplest and most useful presently available. There exists, however, an extensive literature dealing with specialized Lyapunov functions applicable to particular systems. Even an abbreviated discussion of all of these techniques is beyond the scope of the present study. A few are nevertheless worthy of passing notice. The Zubov method(31) is an elegant means of constructing Lyapunov functions; this method is, however, dependent on solving certain partial differential equations whose complexity depends on the system being analyzed. It is therefore useful only for simple systems of low order. Puri and Weygandt(32) develop a method based on analogy with Routh's canonical form for linear systems. A concept of some merit, due to Reiss and Geiss(33) involves the construction of Lyapunov functions, from an assumed $\dot{V}$, by integrating by parts. A recent study by Brockett(34) develops new stability criteria for linear, stationary systems with a nonlinear gain feedback. These criteria involve only the frequency response of the linear portion and the maximum value of the nonlinear gain.

In this section we will conclude the study of Lyapunov stability theory by considering two cases of special interest. The first of these is a very general result which may be derived with very little effort. We consider the system

$$\begin{aligned} \dot{x} &= f(x) \\ f(o) &= 0 \end{aligned} \tag{126}$$

As a Lyapunov function, we take

$$V = x^T Q x \tag{127}$$

where the components of Q are constants. Then we have

$$\dot{V} = f^T Q x + x^T Q f \tag{128}$$

By carrying out the computation for $\frac{d}{d\alpha}[f(\alpha x)]$ it may be verified that

$$f(x) = \int_0^1 J(\alpha x)\, x\, d\alpha \tag{129}$$

where J(x) is the Jacobian matrix of f(x). Substituting this in (128), we find

$$\dot{V} = \int_0^1 x^T \left[J^T(\alpha x)\, Q + Q J(\alpha x) \right] x\, d\alpha \tag{130}$$

Therefore, by theorem 3 of Section 3.3.2, we may conclude that the system (126) is asymptotically stable in the large if for some positive definite matrix Q, the matrix $[J^T(x) Q + Q J(x)]$ is negative definite for all $x \neq 0$.

As usual, the selection of a suitable matrix Q for the problem at hand is the core of the matter, requiring a judicious blend of experience and ingenuity.

We conclude our discussion of the Lyapunov theory by taking note of some deceptive facets of the stability problem for linear nonstationary systems. It is common practice, in the design of autopilots for launch vehicles, to take so-called "time slices"; in effect, a linear, time-varying system is analyzed by assuming that it is valid to consider the system as non-time-varying during preselected small time intervals. Essentially, the assumption may be stated as follows. If in the system

$$\dot{x} = A(t)\, x \tag{131}$$

the eigenvalues of A(t) have negative real parts for all t, then the system may be presumed stable. That this premise is not true in general may be demonstrated by the following counter-example due to Zubov.[35] For the system described by

$$\begin{aligned} \dot{x}_1 &= (12 \sin 6t \cos 6t - 9 \cos^2 6t - 1)\, x_1 \\ &\quad + (12 \cos^2 6t + 9 \sin 6t \cos 6t)\, x_2 \end{aligned}$$

$$\begin{aligned} \dot{x}_2 &= (9 \sin 6t \cos 6t - 12 \sin^2 6t)\, x_1 \\ &\quad - (12 \sin 6t \cos 6t + 9 \sin^2 6t + 1)\, x_2 \end{aligned}$$

The eigenvalues are -1 and -10 for all t. Yet the solution is

$$x_1 = \alpha_1 \exp(2t)(\cos 6t + 2\sin 6t) + \alpha_2 \exp(-13t)(\sin 6t - 2\cos 6t)$$

$$x_2 = \alpha_1 \exp(2t)(2\cos 6t - \sin 6t) + \alpha_2 \exp(-13t)(2\sin 6t + \cos 6t)$$

which is obviously unstable.

The Lyapunov theory may be applied in straightforward fashion to the system described by (131). Let $C(t)$ be a positive definite matrix for all t. As a Lyapunov function, we take

$$V = x^T C(t)\, x \tag{132}$$

Taking the derivative with respect to time

$$\dot{V} = x^T (A^T C + CA + \dot{C})\, x = x^T Bx \tag{133}$$

To prove stability, one must select a positive definite matrix $C(t)$ such that $\dot{V}$, as given by (133) is negative definite. In certain simple cases, some useful results may be obtained as follows. Let the system of Eq. (131) be of second order with the matrix, $A(t)$, given by

$$A(t) = \begin{bmatrix} 0 & 1 \\ a_1 & a_2 \end{bmatrix}$$

where a_1 and a_2 are functions of t. Assume that the matrix $C(t)$ has the form

$$C(t) = \begin{bmatrix} 1 & 0 \\ 0 & c_{22}(t) \end{bmatrix}$$

It is required that $c_{22}(t) > 0$ for all t. It follows that

$$\dot{V} = 2(1 + a_1 c_{22})\, x_1 x_2 + (2 a_2 c_{22} + \dot{c}_{22})\, x_2^2$$

This expression will be negative semidefinite if

$$1 + a_1 c_{22} = 0 \tag{134}$$

and

$$2 a_2 c_{22} + \dot{c}_{22} < 0 \tag{135}$$

From Eq. (134) we obtain

$$c_{22} = - \frac{1}{a_1} \tag{136}$$

Substituting this in (135) yields

$$- \frac{2 a_2}{a_1} + \frac{\dot{a}_1}{a_1^2} < 0 \tag{137a}$$

Furthermore, the requirement that c_{22} (t) be positive for all t means that

$$- \frac{1}{a_1} > 0 \tag{137b}$$

From the stipulations imposed on C(t), we have

$$\lim_{\|x\| \to \infty} [V(x)] \to \infty \tag{138}$$

We will now show that $\dot{V} \not\equiv 0$ along any solution of the given system other than the origin. In fact

$$\dot{V} = \left(- \frac{2 a_2}{a_1} + \frac{a_1}{a_1^2} \right) x_2^2$$

which is zero only along the x_1 axis. But the slope of the system trajectory is

$$\frac{d x_2}{d x_1} = a_1 \frac{x_1}{x_2} + a_2$$

which is infinite for $x_2 = 0$. Therefore the x_1 axis cannot be a trajectory of the system. This proves the assertion. Invoking theorem 4 of Section 3.3.2, we may conclude that if the inequalities

$$- \frac{1}{a_1} > 0 \tag{139}$$

and

$$- \frac{2 a_2}{a_1} + \frac{\dot{a}_1}{a_1^2} < 0 \tag{140}$$

are satisfied, then the system is asymptotically stable in the large.

A result of considerable interest for launch vehicle control systems can be derived from the above analysis. We take a simplified version of an autopilot-controlled space launch vehicle, whose geometry is depicted in Fig. 58, with the control loop schematic as shown in Fig. 59.

The symbols have the following meaning:

I	=	vehicle moment of inertia, slug-ft^2
K_A	=	servo amp. gain, N. D.
K_R	=	rate gyro gain, sec
ℓ_c	=	thrust moment arm, ft
ℓ_α	=	aerodynamic moment arm, ft
L_α	=	aerodynamic load, lb/rad
s	=	Laplace operator, sec^{-1}
T_c	=	thrust, lb
δ	=	thrust angle
θ	=	pitch angle
μ_c	=	$T_c \ell_c / I$
μ_α	=	$L_\alpha \ell_\alpha / I$

All of the quantities T_c, L_α, I, ℓ_c, and ℓ_α are known functions of time. From Fig. 59, we find (writing $x_1 = \theta$)

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = a_1 (t) x_1 + a_2 (t) x_2$$

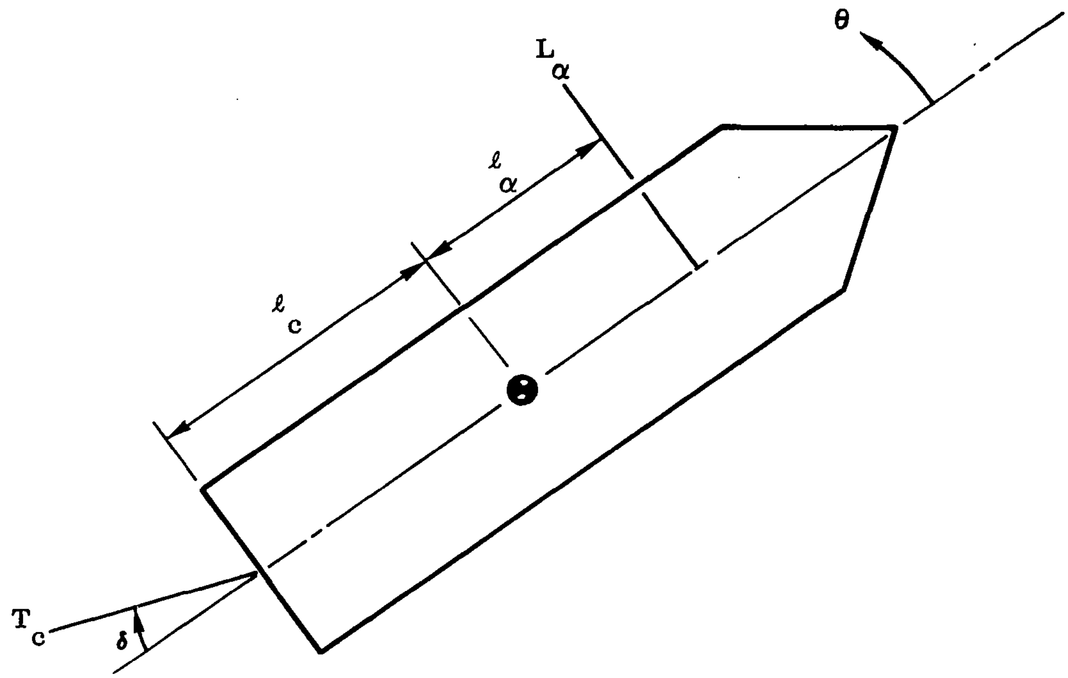


Figure 58. Geometry for Simplified Attitude Control System

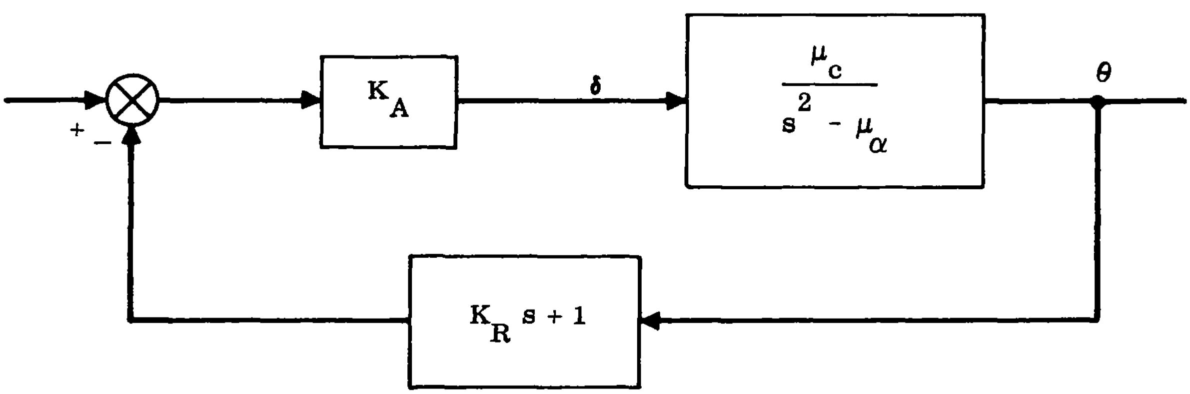


Figure 59. Control Loop for Simplified Attitude Control System

where

$$a_1(t) = \mu_\alpha(t) - K_A \mu_c(t)$$

$$a_2(t) = - K_A K_R \mu_c(t)$$

The foregoing analysis is directly applicable. We find, in fact, that the stability requirements are:

$$K_A \mu_c - \mu_\alpha > 0$$

$$2 K_A K_R \mu_c > \frac{\dot{\mu}_\alpha - K_A \dot{\mu}_c}{(K_A \mu_c - \mu_\alpha)}$$

The first of these is the usual condition derived from linear stationary analysis. The second condition displays quantitatively the destabilizing influence of rapid increase in μ_α.

3.4 ON-OFF CONTROLLERS

One important class of nonlinear control systems is characterized by the fact that the nonlinear element switches discontinuously between extremes. These systems are variously called on-off, bang-bang, or relay type. They find wide application in such areas as roll control for boosters, attitude control for satellites and re-entry vehicles, etc.

The use of phase plane methods is appropriate for the analysis of these systems, and the techniques of Section 3.1 are directly applicable. However, there are many properties of on-off controllers which merit their study as a separate topic. This will be done in the present section.

A very general type of on-off control system is shown in Fig. 60. The properties of the relay are shown in Fig. 61. We will proceed to the analysis of this system from its simplest form, adding various complexities progressively and finally giving a solution for the complete system. In this way, the manner in which each of the system parameters (such as lead filter, time delay, dead zone, and hysteresis) affect the total performance will be made evident.

3.4.1 Ideal Relay

The simplest form of the system of Fig. 60 is with no lead filter or time delay and with no dead zone or hysteresis in the relay. In other words,

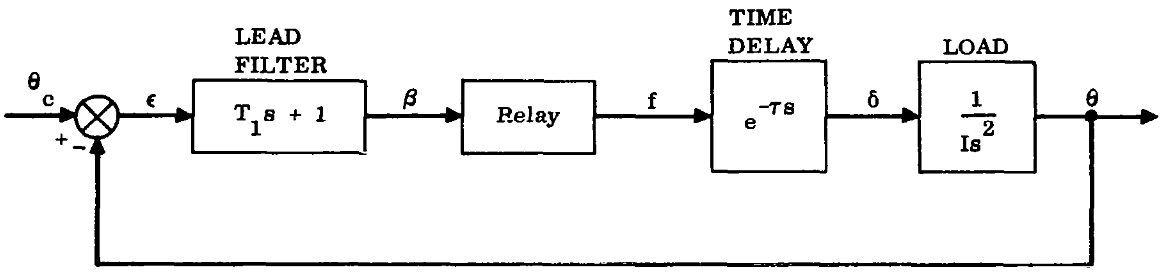


Figure 60. A General Form of On-Off Control System

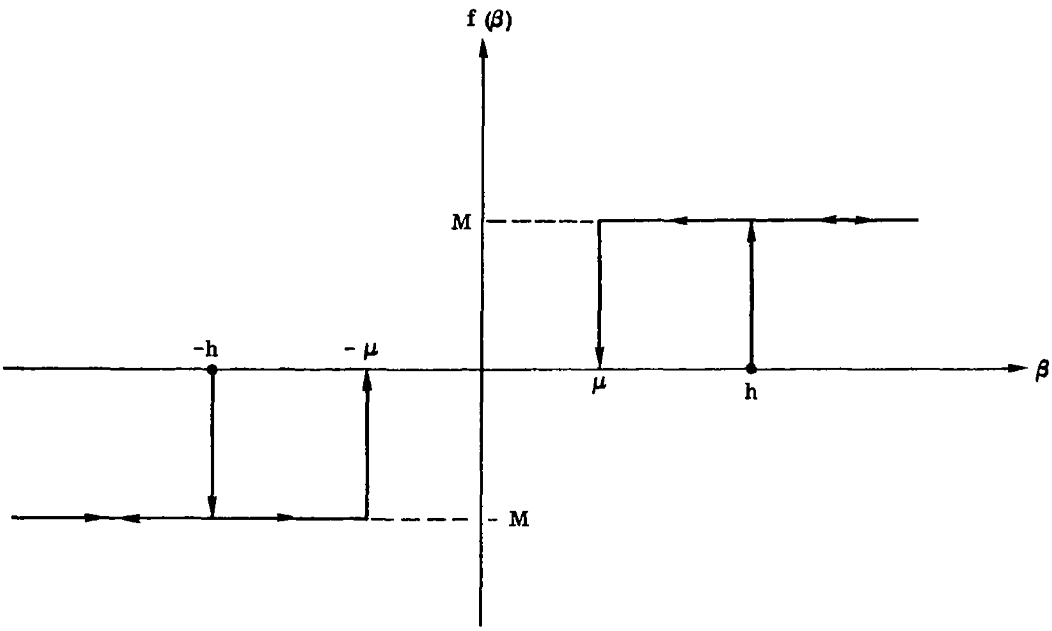


Figure 61. Relay Characteristics

$T_1 = \tau = h = \mu = 0$. We will also assume that there is no input; i.e., $\theta_c = 0$. The system equation then takes the form*

$$I\ddot{\epsilon} + M \operatorname{sgn} \epsilon = 0 \tag{141}$$

and the initial conditions are

$$\epsilon(o) = \epsilon_o$$

$$\dot{\epsilon}(o) = \dot{\epsilon}_o$$

When $\epsilon > 0$, Eq. (141) becomes

$$\ddot{\epsilon} = -\frac{M}{I} \tag{142}$$

Integrating once, we obtain

$$\dot{\epsilon} = -\frac{M}{I}(t - a_1) \tag{143}$$

where

$$a_1 = \frac{I}{M}\dot{\epsilon}_o$$

Performing a second integration yields

$$\epsilon - a_2 = -\frac{M}{2I}(t - a_1)^2 \tag{144}$$

where

$$a_2 = \epsilon_o + \frac{I\dot{\epsilon}_o^2}{2M}$$

Eliminating $(t-a_1)$ between Eqs. (143 and (144) results in

$$\dot{\epsilon}^2 = -\frac{2M}{I}(\epsilon - a_2) \tag{145}$$

$$* \operatorname{sgn} \epsilon = 1 \quad \text{if } \epsilon > 0$$
$$= -1 \quad \text{if } \epsilon < 0$$

This equation represents the system trajectory in the phase plane. It is a parabola with vertex at $\epsilon = a_2$, $\dot{\epsilon} = 0$, and focus at $\epsilon = a_2 - \frac{2M}{I}$, $\dot{\epsilon} = 0$. When $\epsilon < 0$, the motion is governed by

$$\ddot{\epsilon} = \frac{M}{I}$$

Proceeding in the same manner as before, we find for the equation of the system trajectory in this case

$$\dot{\epsilon}^2 = \frac{2M}{I} (\epsilon - a_2') \tag{146}$$

where

$$a_2' = \epsilon_o - \frac{I \dot{\epsilon}_o^2}{2M}$$

Eq. (146) represents a parabola similar to Eq. (145) but of opposite concavity.

The complete trajectory of the system appears as shown in Fig. 62. It is merely a limit cycle whose amplitude depends on the initial conditions.

3.4.2 Relay with Dead Zone

If the relay has a dead zone of magnitude 2μ, then the trajectory (145) is valid as long as $\beta > \mu$. When $-\mu < \beta < \mu$, $M = 0$, and the trajectory in the phase plane is described by $\dot{\epsilon} = \dot{\epsilon}_1$, where $\dot{\epsilon}_1$ is the value of $\dot{\epsilon}$ at point 1 in Fig. 63. Since at this time $\epsilon = \dot{\epsilon}_1 t + \epsilon_1$, it is inferred that ϵ increases for positive $\dot{\epsilon}_1$ and vice versa. There is now no difficulty in completing the phase plane representation for this case as shown in Fig. 63. Again the magnitude of the limit cycle depends on the initial conditions.

3.4.3 Ideal Relay with Lead Circuit

In this case, $\mu = h = \tau = 0$, and T_1 is finite. We have

$$f = M \operatorname{sgn} (T_1 \dot{\epsilon} + \epsilon)$$

Consequently, when $(T_1 \dot{\epsilon} + \epsilon)$ is positive, the motion follows Eq. (145). M changes sign when

$$T_1 \dot{\epsilon} + \epsilon = 0 \tag{147}$$

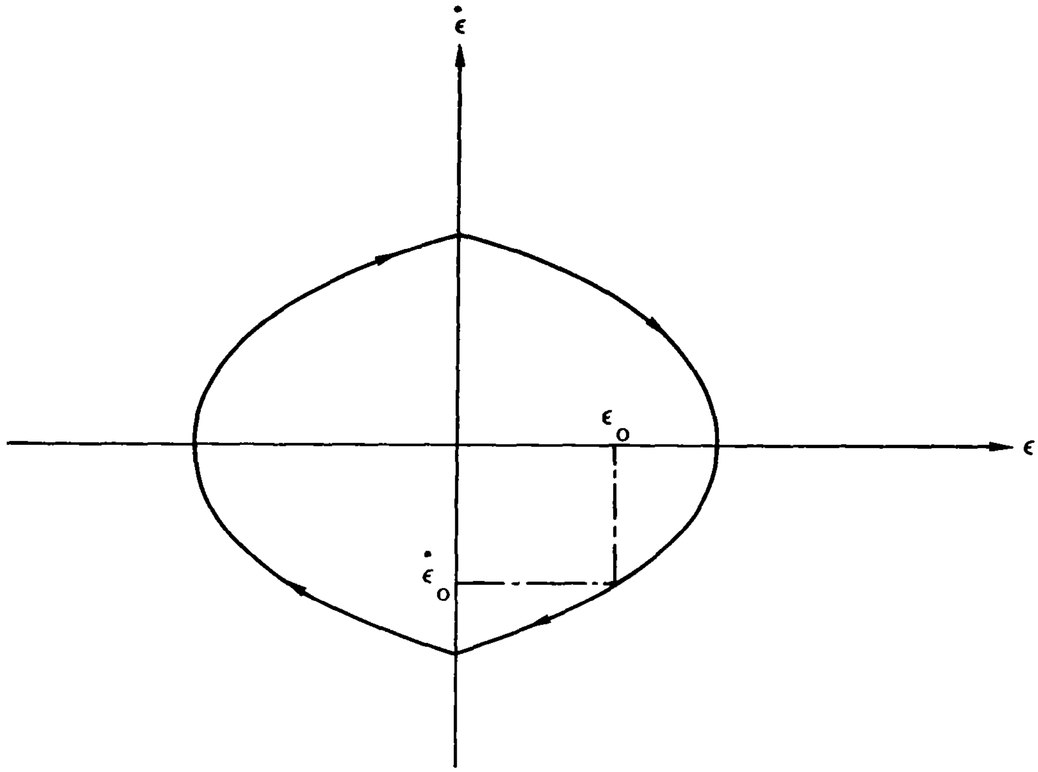


Figure 62. Phase Portrait for Ideal Relay Control System

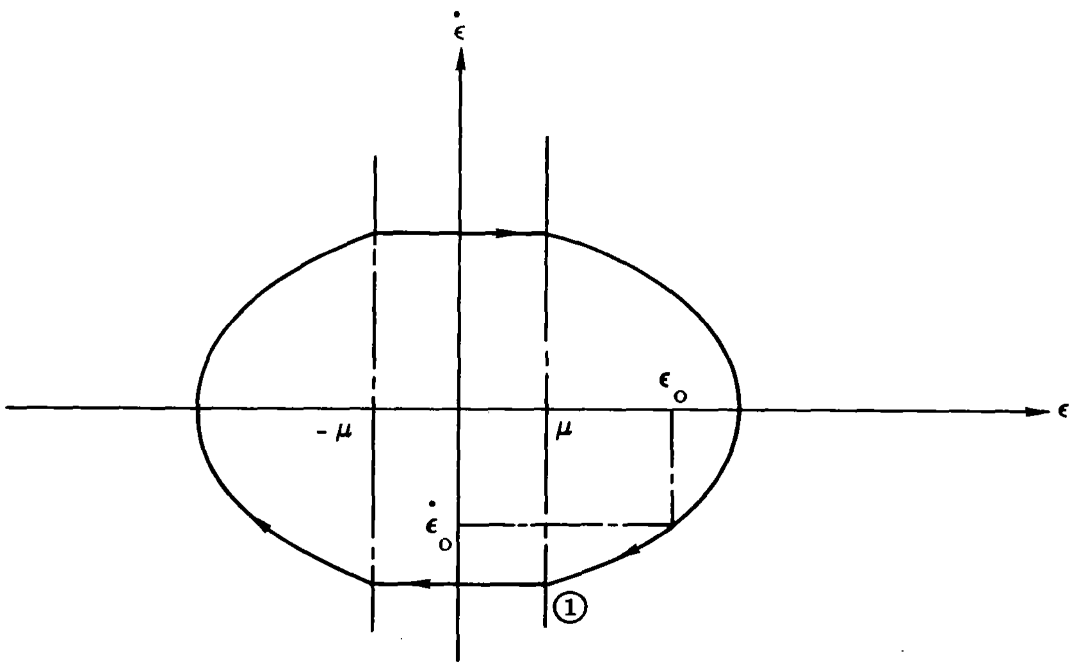


Figure 63. Phase Portrait Showing Influence of Dead Zone

after which the motion proceeds as per Eq. (146). In this way, the phase plane trajectory of Fig. 64 is constructed. When the trajectory reaches point (3) in this diagram, an ambiguity arises. The motion tends to continue in that portion of the phase plane to the right of the switching line. However the tangent to the trajectory parabola has a slope whose magnitude is less than that of the switching line. The parabola representing motion to the right of the switching line appears to the left of the switching line. This is shown by the dotted portion in Fig. 64. There is apparently no way to describe the motion beyond this point (which is obviously not a point of equilibrium). This paradox is, in fact, a result of the assumption of a perfect relay, where it is implicitly assumed that changes of state may take place at infinite speed. By introducing a small time lag, it may be shown that the trajectory converges to the origin exhibiting the well known high frequency, low amplitude oscillation (chatter). This will be done in the next section.

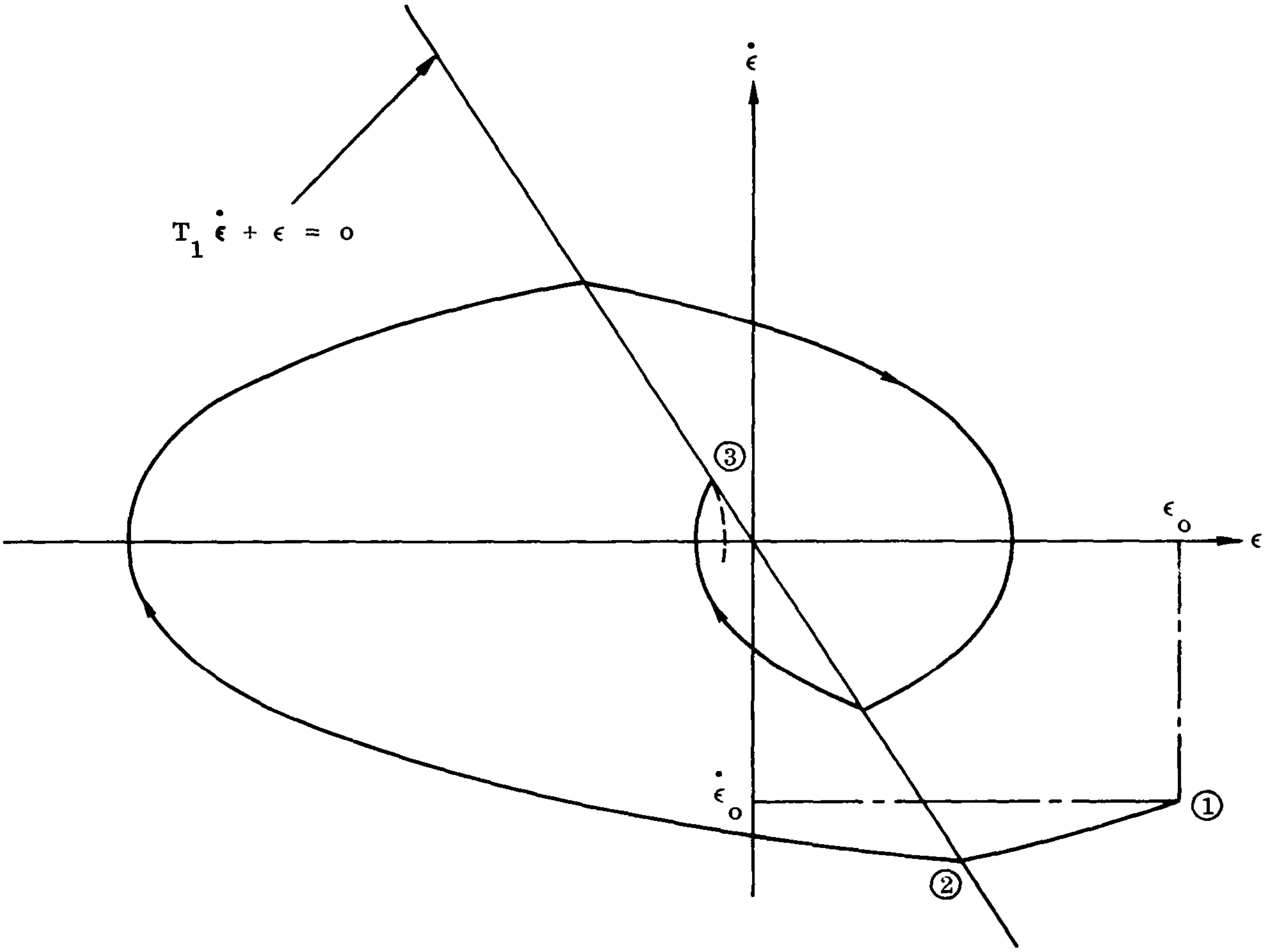


Figure 64. Phase Portrait Showing Effect of Lead Circuit

3.4.4 Relay with Lead Circuit and Time Delay

The situation to be considered here is the same as that in the previous section but with the addition of a finite time delay, τ. To construct the phase plane portrait for this case, we need to develop the equations of the "delayed" switching lines.

Eq. (147) gives the ideal switching line when $\tau = 0$. However, when τ is finite, actual switching occurs at some later time. Let time, t_1, denote the time corresponding to ideal switching and write

$$\epsilon_1 = \epsilon(t_1) \tag{148}$$

$$\dot{\epsilon}_1 = \dot{\epsilon}(t_1) \tag{149}$$

We may determine t_1 as follows. From Eqs. (144) and (143)*

$$\epsilon_1 = -\frac{M}{2I} t_1^2 + \epsilon_o \tag{150}$$

$$\dot{\epsilon}_1 = -\frac{M}{I} t_1 \tag{151}$$

Since these values of ϵ_1 and $\dot{\epsilon}_1$ must satisfy (147), we find

$$t_1 = -T_1 + \sqrt{T_1^2 + \frac{2I}{M}\epsilon_o} \tag{152}$$

The positive sign is used in front of the radical since t_1 is positive. Now the time at which actual switching occurs is given by

$$t_2 = t_1 + \tau \tag{153}$$

We write

$$\epsilon_2 = \epsilon(t_2)$$

$$\dot{\epsilon}_2 = \dot{\epsilon}(t_2)$$

* For simplicity we have assumed that $\dot{\epsilon}_o = 0$.

Substituting in Eq. (150)

$$\epsilon_2 = -\frac{M}{2I} t_2^2 + \epsilon_o$$

$$= -\frac{M}{2I}(t_1 + \tau)^2 + \epsilon_o$$

Using the value of t_1 given by Eq. (152), we find

$$\frac{I\left(\epsilon_2 + \frac{M\tau^2}{2I}\right)}{M(T_1 - \tau)} + T_1 = \left(T_1^2 + \frac{2I}{M}\epsilon_o\right)^{\frac{1}{2}} \tag{154}$$

From Eq. (151)

$$\dot{\epsilon}_2 = -\frac{M}{I}(t_1 + \tau)$$

$$= -\frac{M}{I}\left[-T_1 + \left(T_1^2 + \frac{2I}{M}\epsilon_o\right)^{\frac{1}{2}} + \tau\right]$$

Using Eq. (154) this becomes

$$\dot{\epsilon}_2 = \frac{1}{(T_1 - \tau)}\left[\epsilon_2 + \frac{M\tau}{2I}(2T_1 - \tau)\right] \tag{155}$$

Eq. (155) represents the delayed switching line when M changes sign from positive to negative. To obtain the equation for the delayed switching line when M changes sign from negative to positive, we proceed in identical fashion using Eq. (146) instead of Eq. (145) to obtain an equation analogous to Eq. (150). The procedure is otherwise the same. We obtain finally

$$\dot{\epsilon}_2 = \frac{1}{(T_1 - \tau)}\left[\epsilon_2 - \frac{M\tau}{2I}(2T_1 - \tau)\right] \tag{156}$$

Fig. 65 depicts the character of the phase plane trajectory for this case. We have indicated by ℓ_1 and ℓ_2 the delayed switching lines corresponding to Eqs. (155) and (156) respectively. The case shown is typical for the condition, $T_1 \gg \tau$. When τ is on the same order of magnitude as T_1, a limit cycle of the form shown in Fig. 66 is produced.

3.4.5 Relay with Dead Zone and Hysteresis

The characteristics of the relay were shown in Fig. 61, and we assume that the lead time constant is finite. For the present we assume that $\tau = 0$.

Let $(T_1 \dot{\epsilon} + \epsilon)$ be positive and decreasing. Then the motion follows Eq. (144). When

$$T_1 \dot{\epsilon} + \epsilon = \mu \tag{157}$$

f switches to zero and the trajectory is described by $\dot{\epsilon} = \dot{\epsilon}_1$, where t_1 is the time at which the switching occurs. When

$$T_1 \dot{\epsilon} + \epsilon = -h \tag{158}$$

f switches to -M. Continuing in this way, we construct the phase plane portrait of Fig. 67. Here, lines ℓ_1 and ℓ_2 represent the switching lines (157) and (158) respectively.

Using the approach developed in the previous section it is now a straightforward procedure to include the effect of finite delay, τ. Fig. 68 shows the construction of the phase plane portrait, and Fig. 69 depicts the terminal limit cycle for this case.

3.5 THE POPOV METHOD

For many years, most of the research studies in the area of nonlinear controls have been concerned with extending the basic ideas of the Lyapunov theory and fitting it to the needs of modern automatic control systems. In Section 3.3, we summarized some of the salient features of the current state of the theory. As already noted, one of the chief problems in practical application is the discouragingly complex algebraic manipulations required when the order of the system being analyzed is moderately high.

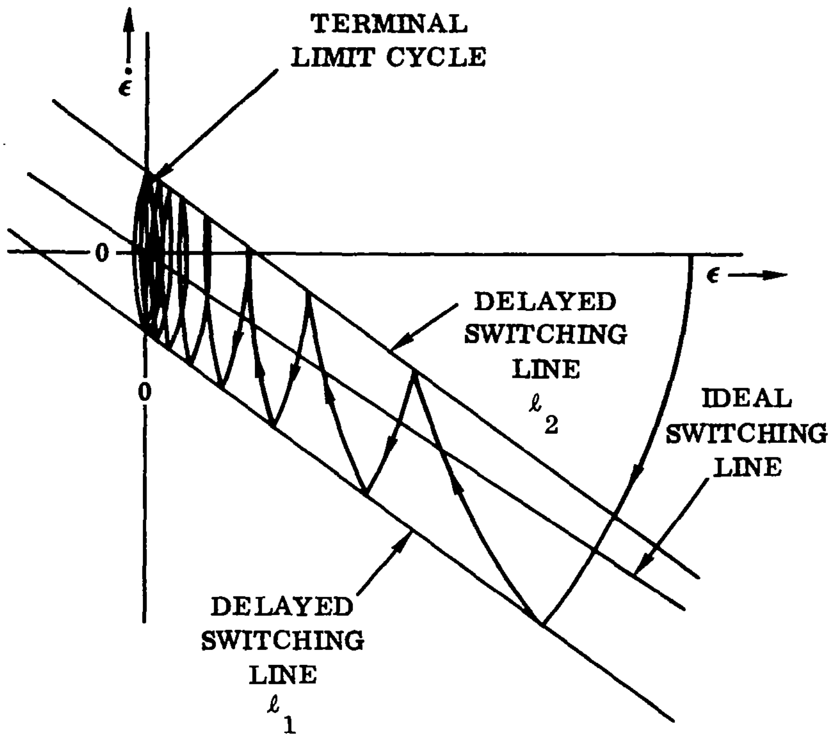


Figure 65. Limit Cycle Produced by Small Time Lag

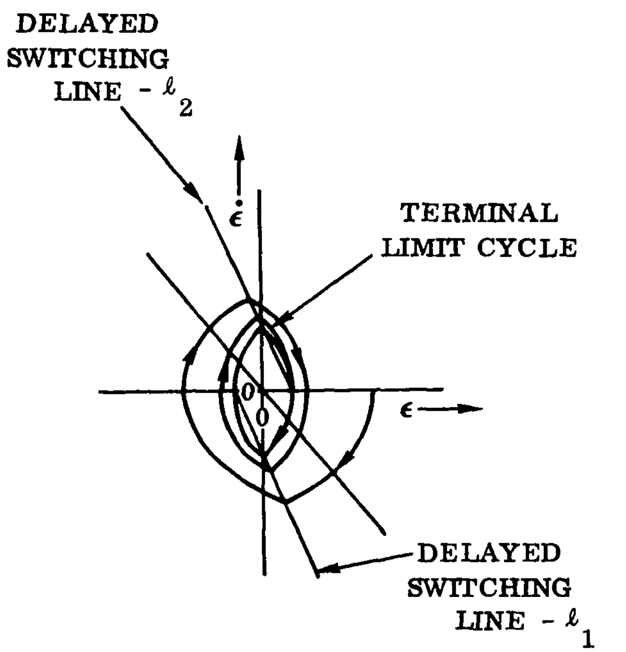


Figure 66. Limit Cycle Produced by Large Time Lag

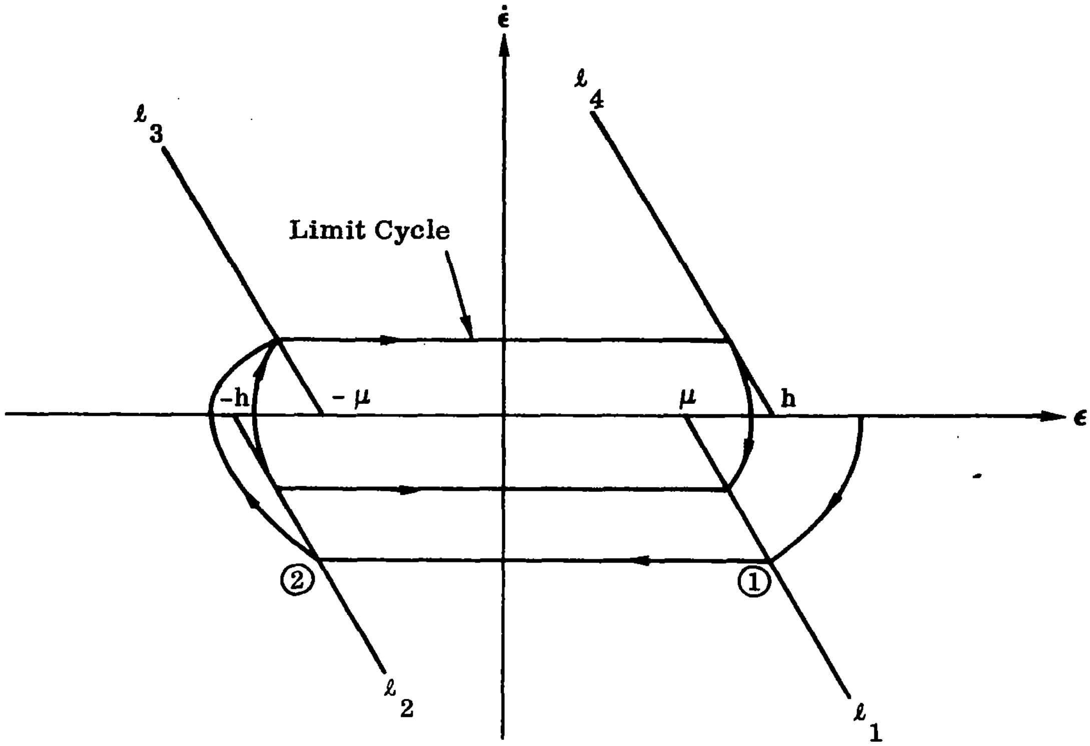


Figure 67. Phase Portrait for Relay with Dead Zone and Hysteresis

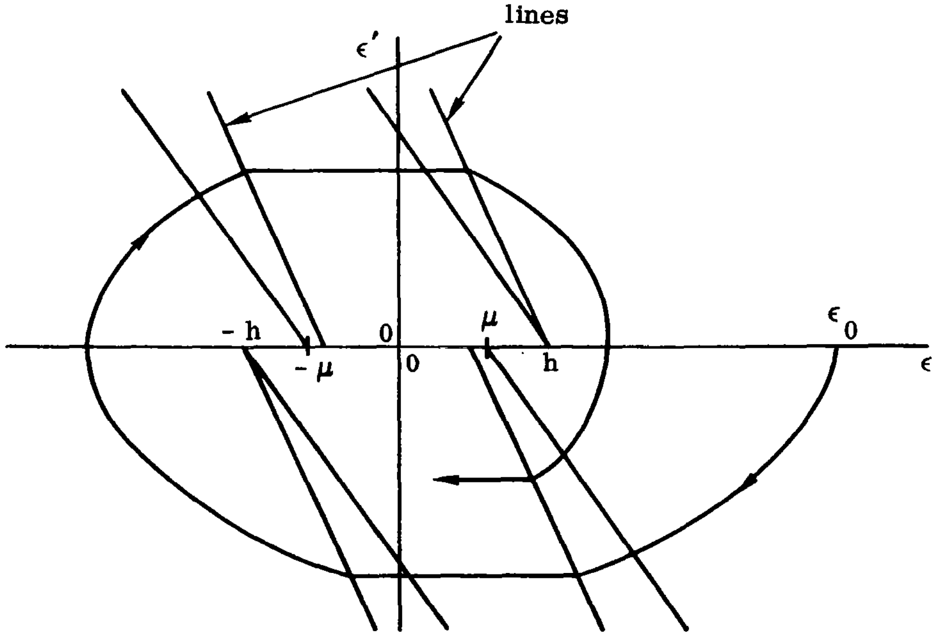


Figure 68. Construction of Phase Portrait for Relay with Dead Zone, Hysteresis and Pure Time Lag

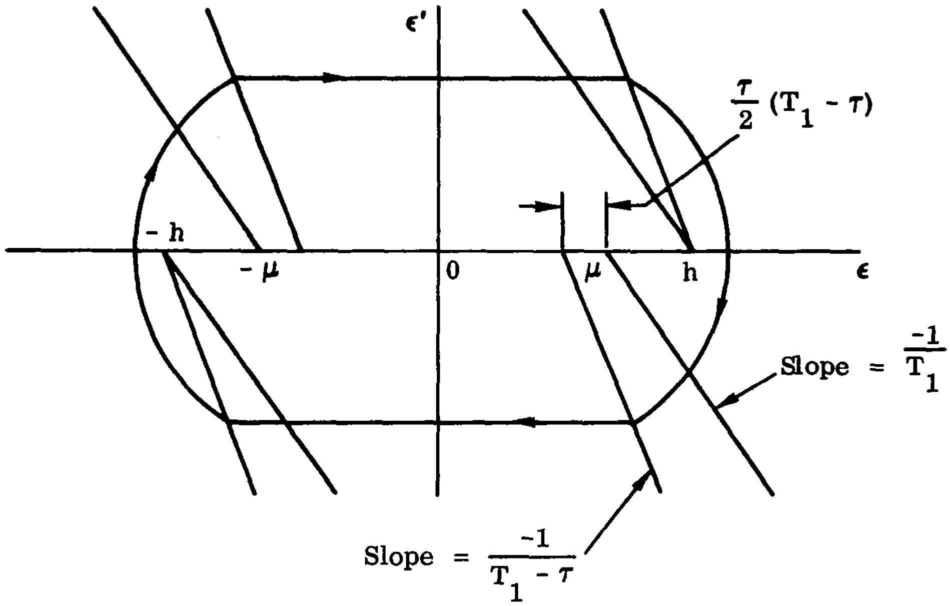


Figure 69. Final Limit Cycle Resulting from Construction of Figure 68

In what is apparently one of the first papers to leave the Lyapunov "fold" in studying nonlinear systems, Popov[36] derived stability criteria in terms of frequency response. He considered systems which contain a linear part and one nonlinear element subject to rather mild restrictions. His results are otherwise very general.

The chief virtue in the stability criteria developed by Popov is that useful results are easy to obtain even for high order systems. Using a computer, one may analyze systems of up to twentieth order -- much in the same way as linear systems.

In general, the systems which may be analyzed using the Popov method encompass a much broader spectrum than those in the Lur'e method. While they are not as completely general as the Lyapunov theory, they are nevertheless of much greater utility in the areas where they apply.

The system to be considered is described by Eqs. (92) which are repeated here for reference.

$$\dot{x} = Ax + b\varphi(\sigma) \tag{159}$$

$$\sigma = v^T x \tag{160}$$

It is assumed that all the eigenvalues of the constant matrix, A, have negative real parts. This restriction will be relaxed in later sections. Furthermore, $\varphi(\sigma)$ is a continuous scalar function satisfying the condition

$$0 \leqslant \frac{\varphi(\sigma)}{\sigma} \leqslant k \tag{161}$$

In what follows, we shall need an explicit expression for the transfer function of the linear portions of Eqs. (159) and (160). This may be derived as follows. Taking the Laplace transform of Eq. (159) yields

$$s\, x(s) = Ax(s) + b\varphi(s)$$

or

$$x(s) = (Is - A)^{-1}\, b\varphi(s)$$

Using Eq. (160), $\sigma = v^T x$

$$G(s) \equiv -\frac{\sigma(s)}{\varphi(s)} = -v^T (Is - A)^{-1}\, b \tag{162}$$

Popov's result is stated in the following theorem.

Theorem: If, in the system described by Eqs. (159) and (160), the following conditions are satisfied:

a. All the eigenvalues of A have negative real parts.

b. $0 \leqslant \frac{\varphi(\sigma)}{\sigma} \leqslant k$

c. There exists a nonnegative real number q such that*

$$\mathrm{Re}\left[(1 + jq\,\omega)\, G(j\,\omega)\right] + \frac{1}{k} > 0 \tag{163}$$

for all $\omega > 0$, where G is defined by Eq. (162).

Then the origin, $x = 0$, is asymptotically stable in the large.**

*The symbol, Re [ ] denotes "real part of"
**Also called complete or absolute stability.

The inequality (163) may be expressed in several simplified forms. Since the linear transfer function may be written in general as the ratio of two polynomials in s; viz.

$$G(s) = \frac{P(s)}{Q(s)}$$

we have

$$G(j\omega) = \frac{P_1(\omega) + j P_2(\omega)}{Q_1(\omega) + j Q_2(\omega)}$$

where the subscripts 1 and 2 are used to denote real and imaginary parts, respectively. This reduces to

$$G(j\omega) = G_1(\omega) + j G_2(\omega) \tag{164}$$

where

$$G_1 = \frac{P_1 Q_1 + P_2 Q_2}{Q_1^2 + Q_2^2} \tag{165}$$

$$G_2 = \frac{P_2 Q_1 - P_1 Q_2}{Q_1^2 + Q_2^2} \tag{166}$$

As a result, the inequality (163) becomes

$$G_1(\omega) - q\omega\, G_2(\omega) + \frac{1}{k} > 0 \tag{167}$$

Further manipulations are possible in order to analyze complex high order systems in a rather routine fashion. A discussion of these will be deferred in order to show the application of the results presented thus far in two simple examples.

Example 16: We consider a second order version of (159) and (160) with

$$A = \begin{bmatrix} 0 & 1 \\ -10 & -7 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$$

$$v = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

A direct application of Eq. (162) yields

$$G(s) = \frac{-\sigma(s)}{\varphi(s)} = \frac{s + 3}{s^2 + 7s + 10}$$

Fig. 70 is a schematic of this system. We seek to establish an upper limit on k in

$$0 \leqslant \frac{\varphi(\sigma)}{\sigma} \leqslant k$$

which ensures complete stability of the system. From Eqs. (165) and (166) we find

$$G_1 = \frac{4\omega^2 + 30}{\omega^4 + 29\omega^2 + 100}$$

$$G_2 = \frac{-\omega(\omega^2 + 11)}{\omega^4 + 29\omega^2 + 100}$$

Substituting in (167)

$$4\omega^2 + 30 + q\omega^2(\omega^2 + 11) + \frac{1}{k}(\omega^4 + 29\omega^2 + 100) > 0$$

We see that this inequality is satisfied for all $\omega > 0$, for any nonnegative q, and $0 < k < \infty$. Consequently this system is completely stable as long as the nonlinear function is restricted to the first and third quadrants.

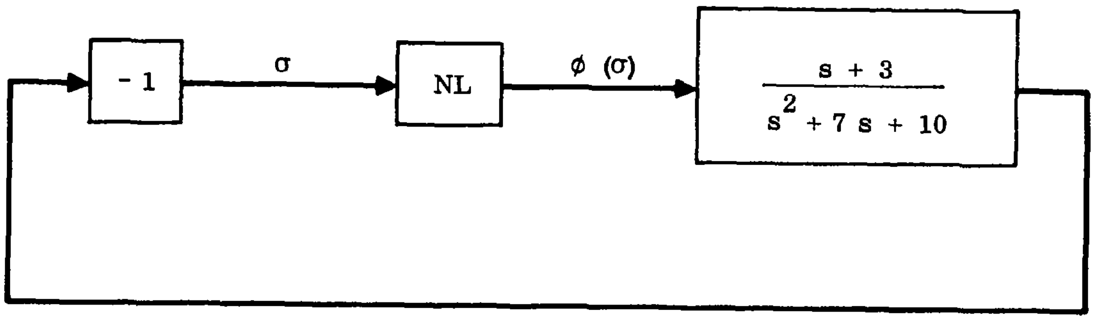


Figure 70. Control Loop for Example 16

Example 17: Consider a third order version of (159) and (160) where

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -20 & -29 & -10 \end{bmatrix}$$

$$b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$v = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$

For third and higher order systems, the matrix inversion to be performed in Eq. (162) becomes extremely awkward, and for all practical purposes impossible, for only moderately high order systems. The essential difficulty stems from the need to carry polynomials in s through the maze of the inversion process. We may circumvent this predicament by first expressing the given system in canonical form in the manner described in Section 3.3.3. Thus by writing

$$x = Ty$$

where T is the Vandermonde matrix given by Eq. (95), we have in the present case

$$\lambda_1 = -1$$

$$\lambda_2 = -4$$

$$\lambda_3 = -5$$

$$T = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -4 & -5 \\ 1 & 16 & 25 \end{bmatrix}$$

$$T^{-1} = -\frac{1}{12}\begin{bmatrix} -20 & -9 & -1 \\ 20 & 24 & 4 \\ -12 & -15 & -3 \end{bmatrix}$$

and we find

$$\dot{y} = Fy + m\varphi$$

$$\sigma = c^T y$$

where

$$F = \begin{bmatrix} -1 & & 0 \\ & -4 & \\ 0 & & -5 \end{bmatrix}$$

$$m = T^{-1} b = \frac{1}{12} \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}$$

$$c = T^T v = - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Since (Is - F) is now a diagonal matrix, the inversion is trivial, and we find from Eq. (162)

$$G(s) = - c^T (Is - F)^{-1} m$$

$$= \frac{1}{(s+1)(s+4)(s+5)}$$

From Eqs. (165) and (166)

$$G_1 = 20 (1 - \omega^2) \left[400 (1 - \omega^2)^2 + \omega^2 (29 - \omega^2)^2\right]^{-1}$$

$$G_2 = - \omega (29 - \omega^2) \left[400 (1 - \omega^2)^2 + \omega^2 (29 - \omega^2)^2\right]^{-1}$$

Substituting in (167), we obtain after some reduction

$$\frac{\omega^6}{k} + \left(\frac{342}{k} - q\right) \omega^4 + \left(\frac{41}{k} + 29 q - 20\right) \omega^2 + \left(\frac{400}{k} + 20\right) > 0$$

As in Example 16, we seek to determine an upper limit on k for which the system is completely stable. This k is obviously dependent on the choice of q. If $q = 0$, then the above inequality reduces to $k < \frac{41}{20}$. A brief inspection of the above inequality indicates however that an optimum upper limit for k may be obtained by taking $q = \frac{342}{k}$, in which case we find $k < 498$.

It is instructive to compare this value with that for the corresponding linear system obtained by replacing $\varphi(\sigma)$ with $\varphi = k\sigma$. This situation is depicted in Fig. 71, with its root locus shown in Fig. 72. Via conventional linear techniques we find that neutral stability occurs for $k = 560$.

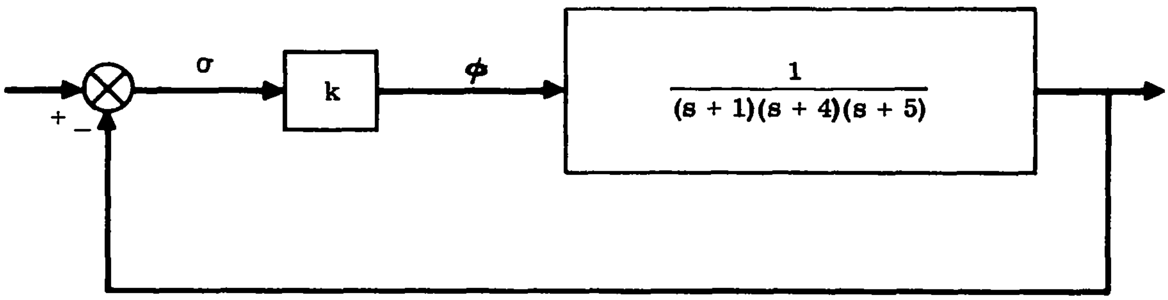


Figure 71. Linearized Version for System of Example 17

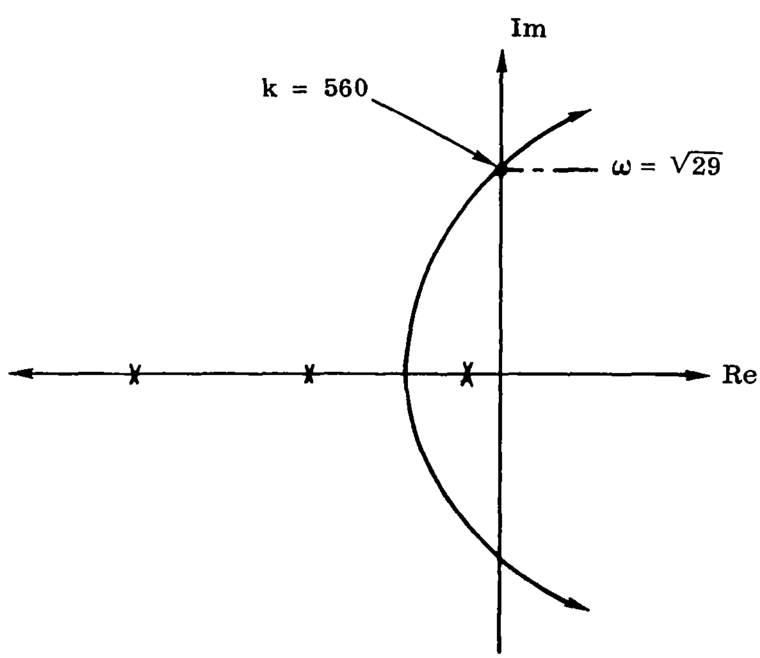


Figure 72. Root Locus for System of Figure 71

In the two examples here considered, it was easy to choose the parameter, q, such that we obtained the maximum upper bound on k for the system to be stable.

For systems of higher order, a somewhat greater effort is necessary to do this but the procedure is still relatively straightforward. To aid in doing this, the inequality (167) is manipulated a little further. We define

$$U = G_1(\omega) \tag{168}$$

$$W = -\omega G_2(\omega) \tag{169}$$

The inequality (167) may therefore be written as

$$U + qW + \frac{1}{k} > 0 \tag{170}$$

$$q \geqslant 0 \quad , \quad k > 0$$

If we replace the inequality in (170) by an equal sign,

$$U + qW + \frac{1}{k} = 0 \tag{171}$$

Then it is easy to show that (170) represents those points in the U - W plane which are to the right of the line (171).

The procedure now consists of making a plot of $G(j\omega)$ in the U - W plane for $0 \leqslant \omega < \infty$. Popov[36] calls this the "modified phase amplitude characteristic" (MPAC). The line (171) is now fitted to this curve such that the curve remains wholly to the right of the line while the abscissa $-\frac{1}{k}$ is made as small as possible. Fig. 73 illustrates the method for a typical $G(j\omega)$ locus.

It is apparent that the usefulness of the method is not seriously compromised by taking systems of arbitrarily high order. The limiting factor is computer time and storage, much the same as for linear system analysis.

In the remaining sections devoted to the Popov method, we will relax various restrictions and discuss some generalizations.

3.5.1 Singular Cases

In the discussion of the Popov method presented in the previous section, it was specifically stipulated that the eigenvalues of the system matrix must have negaative real parts. Those cases in which the system matrix had simple or multiple eigenvalues on the imaginary axis were referred to by Popov as singular. Some of

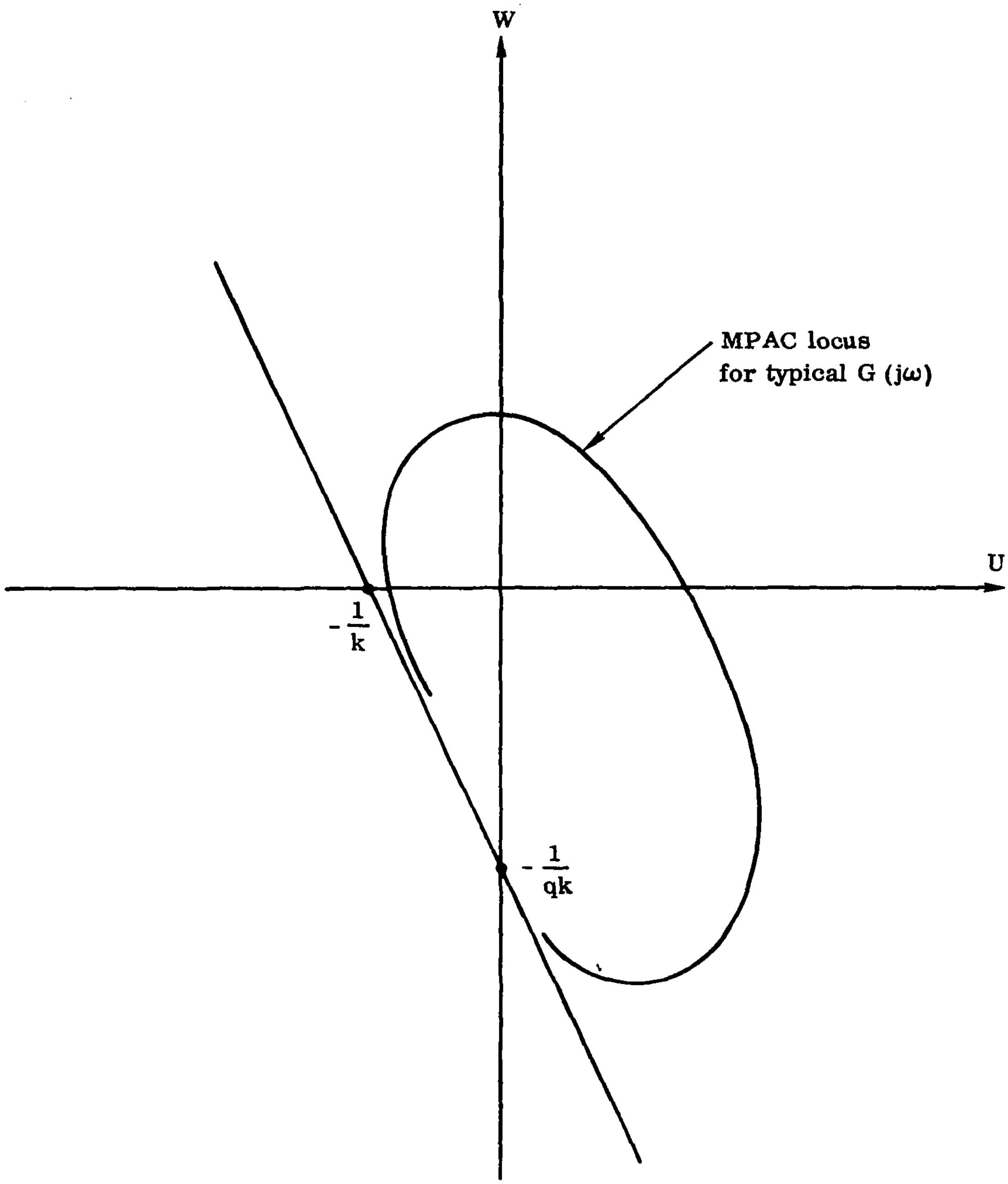


Figure 73. A Typical MPAC Locus Showing Determination of Stability Boundary

these singular cases have been analyzed by Popov[37] and Yakubovitch[41]. However, a simple and elegant result has been obtained by Aizerman and Gantmacher[39] which includes the results obtained by these previous investigators as special cases.

We consider again the system described by Eqs. (159) and (160). The matrix, A, is permitted to have simple or multiple eigenvalues on the imaginary axis. However, instead of (161), the nonlinearity is assumed to satisfy

$$\epsilon \leqslant \frac{\varphi(\sigma)}{\sigma} \leqslant k \tag{172}$$

where ϵ is an arbitrarily small positive number.* The basic result is the following. If

a. Inequality (172) is satisfied.

b. The linear system obtained by substituting $\varphi = \epsilon\sigma$ is stable.

c. There is a real, nonnegative number q such that

$$R_e\,[(1 + j\omega q)\,G(j\omega)] + \frac{1}{k} > 0$$

holds for all $\omega > 0$.

Then the origin $x = 0$ of the system is completely stable.

Further analysis proceeds in a manner completely analogous to that described in the previous section.

3.5.2 Forced Systems

In all the systems considered thus far, there were no external forces applied. As a matter of fact, there are very few results available on the stability of nonlinear systems excited by external forces. Consequently a recent study by Naumov and Tsypkin[40] which adopts the Popov approach in the case where there are externally applied forces is especially significant. The results obtained are also characterized by a greater generality in that the restrictions on the linear portion of the system are greatly relaxed.

---

*Essentially this means that the slope of the nonlinearity is not permitted to be zero at the origin in the $\sigma - \varphi$ plane.

The criteria developed by Naumov and Tsypkin may be presented in a fairly simple manner, and this will be the subject of the present section. The system to be considered is of the form shown in Fig. 74. The basic result is contained in the following theorem.

Theorem: The system of Fig. 74 is completely stable if the following conditions are satisfied.

a. The applied force f(t) is bounded; that is, $\lim_{t \to \infty} f(t) < \infty$.

b. There exists a nonnegative number, r, such that

$$\mathrm{Re}\left[\frac{G(j\omega)}{1 + rG(j\omega)}\right] + \frac{1}{k - r} \geqslant 0 \tag{173}$$

for all $\omega > 0$.

c. $$r + \epsilon \leqslant \frac{d\varphi(\sigma)}{d\sigma} \leqslant k - \epsilon \tag{174}$$

where ϵ is an arbitrarily small positive number, and k is a positive constant which satisfies (173).

Note that the poles of G(s) are not restricted to lie in the left half plane. This constitutes a significant generalization over the results given in the previous two sections.

In the case where all the poles of G(s) do lie in the left half plane, then we may take $r = 0$. A further notable difference between the forced and unforced systems, is that the stability criteria for the former involve the derivative of the nonlinear element rather than the function itself.

For purposes of practical application, the condition (173) may be manipulated to a form in which frequency response techniques may be employed in a familiar manner. Defining a constant, $B = \frac{k}{r}$, we may write (173) as

$$\mathrm{Re}\left[\frac{k\,G(j\omega)}{B + k\,G(j\omega)}\right] + \frac{1}{B - 1} \geqslant 0 \tag{175}$$

$$0 \leqslant \omega < \infty$$

Now let

$$k\,G(j\omega) = M(\omega) + j\,N(\omega) \tag{176}$$

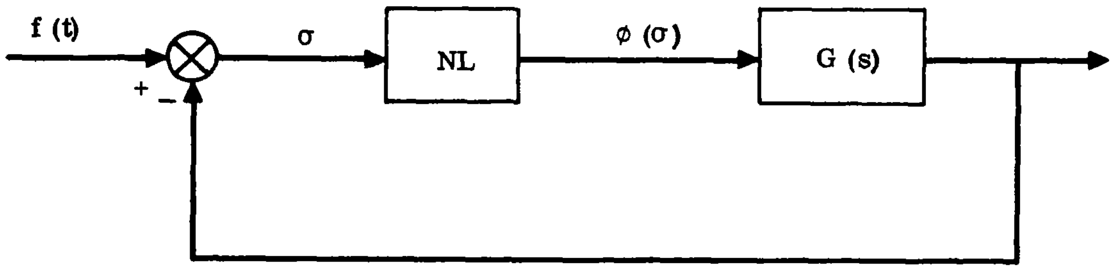


Figure 74. Nonlinear System with Forcing Function

Substituting this in (175) and replacing the inequality by an equal sign, we find

$$[M(\omega) + \frac{1}{2}(B+1)]^2 + N^2(\omega) = \frac{1}{4}(B-1)^2 \tag{177}$$

It is easy to show that this defines a family of circles passing through the point, (-1, j0), having a radius $\frac{1}{2}$ (B - 1), and situated to the left of the line, $M(\omega) = -1$. (See Fig. 75.) Via simple geometric arguments it may be established that inequality (175) is satisfied outside the B circles.

Thus for any given k, if the k $G(j\omega)$ locus in the M - N plane is found to be tangent to a certain B circle, say B_1, this determines a value of r, which in turn defines the limits for $d\varphi(\sigma)/d\sigma$ in order to ensure complete stability. The procedure is suggestive of the use of compensating filters to shape the k $G(j\omega)$ locus in the vicinity of the B circles in order to enhance the stability properties of the system. Conventional linear techniques are appropriate for this purpose.

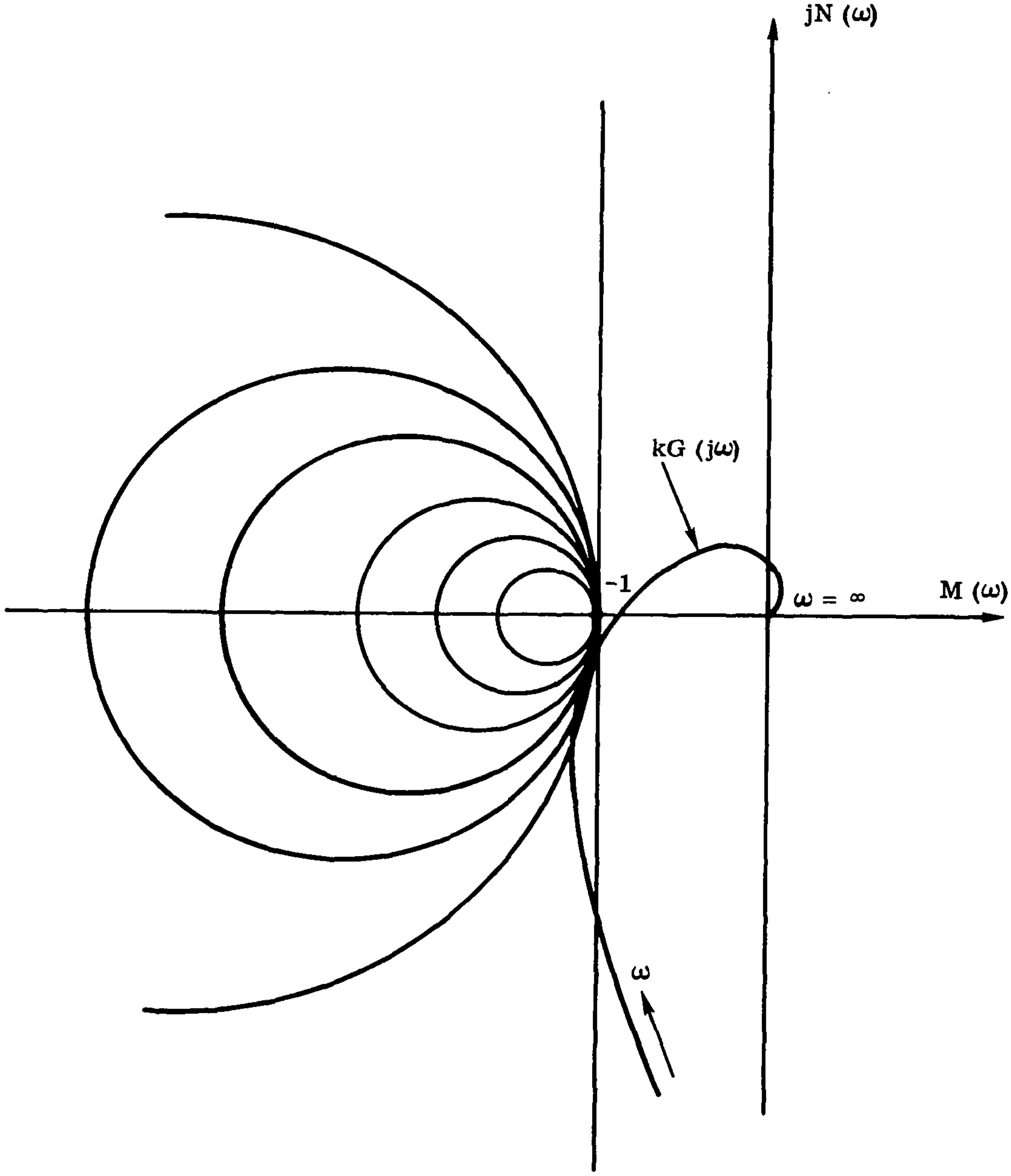


Figure 75. B Circles for Eq. (177)

4. REFERENCES

1. Ku, Y. H., Analysis and Control of Nonlinear Systems, Ronald Press Co., New York, 1958.

2. Grayson, L. P. and Mishkin, E., Three Dimensional Phase Space Analysis, Polytechnic Institute of Brooklyn, Microwave Research Lab., Report R-741-59, PIB-669, May 1959.

3. Kuba, R. E. and Kazda, L. F., "A Phase Space Method for the Synthesis of Nonlinear Servomechanisms," AIEE Trans., Vol. 75, Part II, Nov. 1956, p. 282.

4. Cunningham, W. J., Introduction to Nonlinear Analysis, McGraw Hill, New York, 1958.

5. Minorsky, N., Introduction to Nonlinear Mechanics, J. W. Edwards, Inc., Ann Arbor, Mich., 1947.

6. Davis, H. T., Introduction to Nonlinear Differential and Integral Equations, Dover Publ. Inc., New York, 1962.

7. Kalman, R. E., "Phase Plane Analysis of Automatic Control Systems With Nonlinear Gain Elements," Trans. AIEE, Vol. 73, Part II, 1954, p. 383.

8. Kochenburger, R.J., "Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms," Trans. AIEE, Vol. 69, Part 1, 1950, p. 270.

9. Truxal, J. G., Control System Synthesis, McGraw Hill Co., New York, 1955.

10. Sridhar, R., "A General Method for Deriving the Describing Functions for a Certain Class of Nonlinearities," IRE Trans. on Automatic Control, Vol. AC-5, 1960, p. 135.

11. Kochenburger, R.J., "Limiting in Feedback Control Systems," Trans. AIEE, Vol. 72, Part II, 1953, p. 180.

12. Gran, R. and Rimer, M., "Stability Analysis of Systems with Multiple Nonlinearities," IEEE Trans. on Automatic Control, 1965, p. 94.

13. Jud, H. G., "Limit Cycle Determination for Parallel Linear and Nonlinear Elements," IEEE Trans. on Automatic Control, 1964, p. 183.

14. Halstenberg, R. V., "Combined Hysteresis and Nonlinear Gains in Complex Control Systems," IRE Trans. on Automatic Control, 1958, p. 51.

15. Gronner, A. D., "The Describing Function of Backlash Followed by a Deadzone," AIEE Trans., Applications & Industry, Vol. 77, 1958, p. 403.

16. Gelb, A. and Vander Velde, W. E. "On Limit Cycling Control Systems," IEEE Trans. on Automatic Control, Vol. AC-8, 1963, p. 142.

17. Lyapunov, A., "Problème Général de la Stabilité du Mouvement," Annals of Mathematics Study No. 17, Princeton Univ. Press, Princeton, N. J. (This is a French translation of Lyapunov's original paper which was published in Russia in 1902.)

18. Kalman, R. E. and Bertram, J. E., "Control System Analysis via the Second Method of Lyapunov; I. Continuous Time Systems, II. Discrete Time Systems," Trans. ASME, Vol. 82, Series D, 1960, p. 371.

19. LaSalle, J. P. and Lefschetz, Stability by Lyapunov's Direct Method with Applications, Academic Press, New York, 1961.

20. Lur'e, A. I., Certain Nonlinear Problems in the Theory of Automatic Control, Gosteknizdat, 1951 (Russian).

21. Letov, A. M., Stability in Nonlinear Control Systems, Princeton Univ. Press, Princeton, No. J., 1961 (a translation of the 1955 Russian edition).

22. Rekasius, Z. V. and Gibson, J. E., "Stability Analysis of Nonlinear Control Systems by the Second Method of Lyapunov," IRE Trans. on Automatic Control, Vol. AC-7, 1962, p. 3.

23. Schultz, D. G. and Gibson, J. E., "The Variable Gradient Method for Generating Lyapunov Functions," Trans. AIEE, Vol. 81, Part II, 1962, p. 203.

24. Lasalle, J. P., "Some Extensions of Lyapunov's Second Method," IRE Trans., Vol. CT-7, 1960, p. 520.

25. Laning, J. H. and Battin, R. H., Random Processes in Automatic Control, McGraw Hill Co., New York, 1956.

26. Tou, J. T., An Introduction to Modern Control Theory, McGraw Hill Co., New York, 1964.

27. Bellman, R., Introduction to Matrix Analysis, McGraw Hill Co., New York, 1960.

28. Szego, G. P., "A Contribution to Lyapunov's Second Method: Nonlinear Autonomous Systems," Trans. ASME, J. of Basic Eng., 1962, p. 571.

29. Ingwerson, D. R., A Modified Lyapunov Method for Nonlinear Stability Analysis, IRE Trans. on Automatic Control, Vol. AC-6, May 1961.

30. Lass, H., Vector and Tensor Analysis, McGraw Hill Book Co., New York, 1950.

31. Margolis, S. G. and Vogt, W. G., "Control Engineering Applications of Zubov's Construction Procedure for Lyapunov Functions," IEEE Trans. on Automatic Control, 1963, p. 104.

32. Puri, N. and Weygandt, C., "Second Method of Lyapunov and Routh's Canonical Form," J. of Franklin Inst., Vol. 276, 1963, p. 365.

33. Reiss, R. and Geiss, G., "The Construction of Lyapunov Functions," IEEE Trans. on Automatic Controls, 1963, p. 382.

34. Brockett, R. W., "On the Stability of Nonlinear Feedback Systems," Proc. Joint Automatic Control Conf., Stanford, Calif., 1964, p. 288.

35. Zubov, V., Mathematical Methods for the Study of Automatic Control Systems, Permagon Press, 1962.

36. Popov, V. M., "Absolute Stability of Nonlinear Systems of Automatic Control," Automation and Remote Control, Vol. 22, 1961, p. 857.