From Nonlinear System Analysis and Synthesis: Vol. 2 - Techniques and Applications, published by the ASME, Feb. 1981.

Part 4, Chapter 1

APPLICATIONS OF RIGOROUS STABILITY CRITERIA TO NONLINEAR DYNAMICAL SYSTEMS

James H. Taylor
School of Mechanical and Aerospace Engineering
Oklahoma State University
Stillwater, Oklahoma 74078

1. INTRODUCTION

The primary goal of this presentation is to develop and illustrate a philosophy of nonlinear system stability analysis that recognizes the many great strides that have been made in this field in the last few decades, without dwelling unnecessarily on the theoretical foundations and details of the fundamental results that now exist. It is the author's hope that the practicing engineer, whether or not he has a firm grounding in modern systems theory, will find this material to be helpful in solving concrete, "real-world" problems.

The application of rigorous theorems to determine conditions for stability in nonlinear systems is an area in systems theory that seems to have been somewhat neglected. This anomaly appears to have three major causes. The first is a rather extreme case of the usual gap between theoreticians and practicing engineers. The other two causes are more essentially philosophical in nature; we may call them "conservatism" and "benign neglect."

It is a fact that many rigorous stability assessment methods are conservative, in the sense that the specific system under study may really be stable over a broader set of conditions than the stability theorems will allow. For example, if an absolute stability criterion (as outlined in Section 2 below, and developed more fully in [1]) is applied to show that a system containing the nonlinearity $f(\sigma)$ where σ is the nonlinearity input variable is absolutely stable under the gain sector constraint

$$0 \leq \frac{f(\sigma)}{\sigma} \leq \bar{F} = 1 \tag{1}$$

(we will discuss both absolute stability and gain sectors more fully later in this presentation), then it may be true, for the particular nonlinearity under consideration, that taking the upper bound $\bar{F} = 2$, or even 5, may still result in a stable system. The word particular is stressed above, because it may also be true that there is some strangely behaved nonlinearity $f^*(\sigma)$ for which the condition in Eq. (1) is necessary and sufficient, in the sense that for $\bar{F}^* = 1.1$ the nonlinear system is unstable; however, the practitioner is unconcerned with that eventuality, and probably rightfully so.

Finally, the idea of benign neglect, best described as "building the system and adjusting the gains to make it behave acceptably" is obviously a dangerous expedient when dealing with large, sophisticated and costly systems. It may well be that adjusting the gains won't help, and expensive design modifications will have to be made. It also may be true that the nonlinear effect will only rarely destabilize the system, so its importance will only be recognized after several mysterious failures. The fact is, the possible destabilizing nature of nonlinear effects can and should be considered so that costly mistakes can be avoided.

The first problem (the gap) is historical, and can best be overcome by focusing on the most down-to-earth stability criteria available. This shall be done in the outline in Section 2. The other barriers are perhaps best dealt with by providing some worthwhile examples, as in Sections 3 and 4, and by discussing how absolute stability theorems can be used during the systems design phase, Section 5. The remainder of this paper deals with time-varying systems (Section 6), systems with multiple nonlinearities (Section 7), and a final summary of the usefulness of absolute stability criteria (Section 8).

2. STABILITY THEOREMS AND SYSTEM MODELS

Stability theorems can conveniently be divided into two categories: direct and indirect. The best-known direct stability theorem for nonlinear systems is due to Lyapunov; the following statement is from [1].

Theorem 1 (Lyapunov): Consider the continuous-time free dynamic system

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

where $\underline{f}(\underline{0},t) = \underline{0}$ for all t. If a scalar function $V(\underline{x},t)$ is defined for all $\underline{x}$ and t such that $V(\underline{0},t) = \underline{0}$ and it satisfies

(i) $V(\underline{x},t)$ is positive definite, i.e., there exists a continuous nondecreasing function α such that $\alpha(0) = 0$ and

$$0 < \alpha(||\underline{x}||) \leq V(\underline{x},t), \qquad \underline{x} \neq \underline{0}$$

(ii) $V(\underline{x},t)$ is decrescent, i.e., $V(\underline{x},t) < \beta(||\underline{x}||)$ where β is a continuous scalar nondecreasing function and $\beta(0) = 0$,

(iii) $V(\underline{x},t)$ is radially unbounded, i.e., $\alpha(||\underline{x}||) \to \infty$ with $||\underline{x}|| \to \infty$

(iv) $\dot{V}(\underline{x},t) \triangleq \frac{\partial V}{\partial t} + \underline{\nabla V}^T \underline{f}(\underline{x},t) \leq -\gamma(||\underline{x}||) < 0, \ \underline{x} \neq \underline{0}$

where γ is a continuous scalar function such that $\gamma(0) = 0$

then the equilibrium state $\underline{x} = \underline{0}$ is uniformly asymptotically stable in the large and $V(\underline{x},t)$ is called a Lyapunov function for the system. ●

We consciously avoid defining "uniform asymptotic stability in the large" (UASIL) since it is dealt with in sufficient detail in [1]; it is enough to observe that systems that enjoy this property are not subject to pathological behavior, such as unbounded response to bounded inputs (see [1]), which may be encountered if less stringent stability definitions are used (and thus the conditions of Theorem 1 are relaxed correspondingly).

The Lyapunov theorem is the most general nonlinear system stability result available. It is somewhat tantalizing, however, because it is often an heroic task to apply it to specific systems of order greater than three (or even two, in some instances), because one must search for V so that the conditions of Theorem 1 are satisfied. Occasionally (rarely!) a Lyapunov function suggests itself naturally (e.g. V = potential energy + kinetic energy); otherwise a great deal of work and intuition or prior experience will be required to find V. However, it is important to appreciate the generality and rigor of Lyapunov's contribution to nonlinear system theory --- directly or indirectly, Theorem 1 has served as the cornerstone for a great deal of development. For a good summary of the direct application of Theorem A, refer to Kalman and Bertram [2] and LaSalle and Lefschetz [3].

Mention must also be made of the functional analytic approach to nonlinear systems stability theory. Since this approach is even more abstract than the Lyapunov method, it may not be considered to be an applications-oriented technique; the interested reader may refer to Holtzman [4] and Vidyasagar [5]. The main use of functional analysis has been the derivation of stability criteria which are very nearly equivalent to those stated below except for some details of the system model and definition of stability. To discuss these differences violates the spirit of this presentation.

The direct stability analysis methods outlined above have been used to develop very effective stability criteria for nonlinear and time-varying systems. The major contribution, from the view point of the practicing engineer, is that the direct process defined in Theorem 1 can be completely avoided. The conditions for UASIL given by the criteria can be checked; if they are satisfied, then the proofs detailed in Narendra and Taylor [6] provide an assurance that the provisions of Theorem 1 can be met, without the need for any further analysis.

Before the stability criteria can be stated, a specific system model must be established. The simplest formulation is given in the feedback control system configuration shown in Fig. 1, where the usual manipulations have been performed to place the linear plant dynamics in the forward path, as represented by the transfer function W(s), and the nonlinear time-varying relation, $g(\sigma,t)$, is in the feedback path. For the modern control theorist, the equivalent state space model is given by the vector differential equation

$$\begin{aligned} \dot{\underline{x}} &= \underline{\bar{A}}\,\underline{x} + \underline{b}\tau \\ \sigma &= \underline{h}^T\underline{x} + \rho\tau \\ \tau &= -g(\sigma,t) \end{aligned} \tag{2}$$

where σ and τ correspond to the variables shown in Fig. 1. The first two parts of Eq. (2) describe the linear part of the system; in fact, by direct Laplace transformation

$$W(s) \triangleq \frac{L[\sigma]}{L[\tau]} = \rho + \underline{h}^T(s\underline{I} + \underline{A})^{-1}\underline{b} \tag{3}$$

provides the bridge between the two control theoretic model representations.

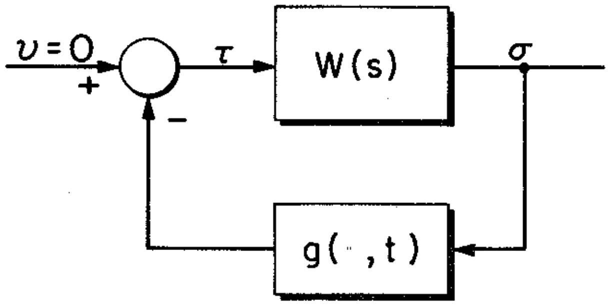


Fig. 1 Basic System Configuration

Finally, for engineers and applied mathematicians without familiarity with the above control systems models, let us illustrate the relation between scalar differential equation models and Fig. 1: Consider

$$\ddot{y} + a_2\dot{y} + a_1 y + g((h_1 y + h_2\dot{y}),t) = 0 \tag{4}$$

First, we can directly define a state vector $\underline{x}$ and re-write Eq. (4) in the form of Eq. (3):

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \triangleq \begin{bmatrix} y \\ \dot{y} \end{bmatrix} \quad \text{yields} \quad \begin{cases} \dot{x}_1 = x_2 \quad \text{(by definition of } x_2) \\ \dot{x}_2 = -a_1x_1 - a_2x_2 - g(\sigma,t); \\ \sigma = h_1x_1 + h_2x_2 \quad \text{(from Eq. (4))} \end{cases}$$

In matrix form, we have

$$\dot{\underline{x}} = \begin{bmatrix} 0 & 1 \\ -a_1 & -a_2 \end{bmatrix} x_2 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tau \quad , \quad \begin{cases} \sigma = [h_1 \quad h_2]\underline{x} = \underline{h}^T\underline{x} \\ \tau = -g(\sigma,t) \end{cases}$$

Then, by applying Eq. (3),

$$W(s) = [h_1 \quad h_2] \begin{bmatrix} s & -1 \\ a_1 & s+a_2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{h_1 + h_2 s}{s^2 + a_2 s + a_1}$$

provides the desired relation between Eq. (4) and Fig. 1.

The above systems models are consistent with [1-3], which was written from the Lyapunov point of view where the main concern is: "If the system is perturbed from its equilibrium $\underline{x} = \underline{0}$, will the solutions return to that equilibrium?"; thus there are <u>no external inputs</u> ($\upsilon \equiv 0$, Fig. 1). The functional analytic approach deals with questions such as "if υ is bounded, then is σ bounded?" This property is called bounded input/bounded output stability (BIBOS). The essential point for our purposes is that the criteria that follow guarantee stability in both senses, so there is no need to be concerned with this distinction, or with the presence of inputs.

Before considering the stability of a nonlinear time-varying system, it is important to know the range of linear feedback relations — i.e. the range of k in the relation

$$\tau = -k\sigma \tag{5}$$

which leads to an asymptotically stable system in Fig. 1 or Eq. (2). The importance of such a range (e.g., assume that $\underline{K} < k < \overline{K}$ guarantees asymptotic stability) is that the nonlinear and/or time-varying characteristic $g(\sigma,t)$ will have to be confined to this range, in the sense that

$$\underline{K} < \underline{G} \leq \frac{g(\sigma,t)}{\sigma} \leq \overline{G} < \overline{K} \tag{6}$$

where $\underline{G}$ and $\overline{G}$ are the actual bounds on g/σ. The significance of this concept of the range of a nonlinearity is illustrated in Fig. 2; generally, we speak of confining $g(\sigma,t)$ to the sector defined by $\underline{G}$ and $\overline{G}$, or Eq. (6) is referred to as a sector constraint. Determining $\underline{K}$, $\overline{K}$ for linear systems is accomplished by using Nyquist's Criterion, the Routh-Hurwitz Conditions, or the root-locus method (all of which are outlined in [1]).

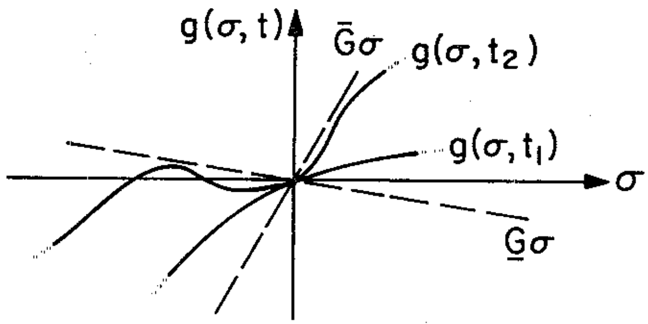


Fig. 2 Gain Sector for a Nonlinear Time-Varying Element

All of the stability criteria discussed below involve frequency domain constraints on $W(j\omega)$, Eq. (3). The conditions are graphical in nature, and are based on one of the following plots:

Nyquist plot: The graph of $V(\omega) \triangleq \text{Im } W(j\omega)$ versus $U(\omega) \triangleq \text{Re } W(j\omega)$

Popov plot: The graph of $\hat{V}(\omega) \triangleq \omega \text{ Im } W(j\omega)$ versus $U(\omega) \triangleq \text{Re } W(j\omega)$

These plots are illustrated in the examples that follow (cf. Figs. 3 and 4).

With the above preliminaries as a background, we restate the absolute stability criteria that will be applied to nonlinear systems in this presentation [1]:

Theorem 2 (Circle Criterion): A system of the form portrayed in Fig. 1 is UASIL (BIBOS) if for any $g(\sigma,t)$ obeying Eq. (6) the Nyquist plot of $W(j\omega)$ does not touch or intersect the circle whose diameter is $-1/k$, $\underline{G} \leq k \leq \overline{G}$. ●

Theorem 3 (Parabola Criterion): A system of the form portrayed in Fig. 1 is UASIL (BIBOS) if for any $f(\sigma)$ obeying

$$\underline{K} < \underline{F} \leq \frac{f(\sigma)}{\sigma} \leq \overline{F} < \overline{K} \tag{7}$$

the Popov plot of $W(j\omega)$ does not touch or intersect the parabola

$$(\underline{F}U + 1)(\overline{F}U + 1) = (\overline{F} - \underline{F})\hat{V}/\alpha \tag{8}$$

for some value of α. ●

Theorem 4 (Off-axis Circle Criterion): A system of the form portrayed in Fig. 1 is UASII (BIBOS) if for any monotonic* $f(\sigma)$ obeying

$$\underline{K} < \underline{M} \leq \frac{df(\sigma)}{d\sigma} \leq \overline{M} < \overline{K} \tag{9}$$

the Nyquist plot of $W(j\omega)$ for $\omega \geq 0$ does not touch or intersect a circle having $-1/k$, $\underline{M} \leq k \leq \overline{M}$ as a chord. ●

3. PRELIMINARY ILLUSTRATIONS AND COMMENTS

The Circle Criterion (CC) is discussed in an introductory fashion in [1]. It is mentioned that for a given $W(s)$, an infinite number of circles can be drawn that avoid $W(j\omega)$ on the Nyquist plot, resulting in an infinite number of sector bounds $\underline{G}$ and $\overline{G}$. The five principal cases are illustrated in Fig. 3; these are

(a) $0 \leq \frac{g(\sigma,t)}{\sigma} \leq \overline{G}_1$ and $0 < \underline{G}_2 \leq \frac{g(\sigma,t)}{\sigma} \leq \overline{G}_2$

(b) $0 > \underline{G}_3 \leq \frac{g(\sigma,t)}{\sigma} \leq \overline{G}_3 > 0$

(c) $\underline{G}_4 \leq \frac{g(\sigma,t)}{\sigma} \leq 0$ and $\underline{G}_5 \leq \frac{g(\sigma,t)}{\sigma} \leq \overline{G}_5 < 0$

The two cases where one bound is zero result in degenerate circles (vertical straight lines), as shown. Also, it should be observed that the circle must enclose the Nyquist plot of $W(j\omega)$ if $\underline{G}$ is negative and $\overline{G}$ is positive (Fig. 3b). Note that the "diameter", $-1/k$ for $\underline{G} < k \leq \overline{G}$, is outside the circle in this case.

The Parabola Criterion (PC) is a bit more difficult to apply, since parabolas are not easy to draw exactly, and you have one degree of freedom in the choice of α. Thus, let us first consider the special case $\underline{F} = 0$, which is the celebrated Popov Criteria [1]. Equation (8) then reduces to

$$\hat{V} = \alpha(U + 1/\overline{F}) \tag{10}$$

so the condition for UASIL (BIBOS) is simply that the Popov plot of $W(j\omega)$ must not touch or intersect a straight line passing through $-1/\overline{F}$ with arbitrary slope α, as illustrated in Fig. 4. The slope of the Popov line may be positive or negative but not zero.

For the general case, the estimation of $\underline{F}$ and $\overline{F}$ such that a suitable parabola can be drawn is simplified by noting that the parabola specified in Eq. (8) has several useful properties:

(i) Real axis crossing ($\hat{V} = 0$) are at $U = -1/\underline{F}$ and $U = -1/\overline{F}$.

(ii) The parabola is tangent to straight lines drawn through these

---

* The term monotonic [1] in this context is a standard generalization of the concept of monotonic increasing functions i.e., $f(\sigma)$ such that $df/d\sigma \geq 0$. It might be clearer to call functions satisfying Eq. (9) "slope constrained."

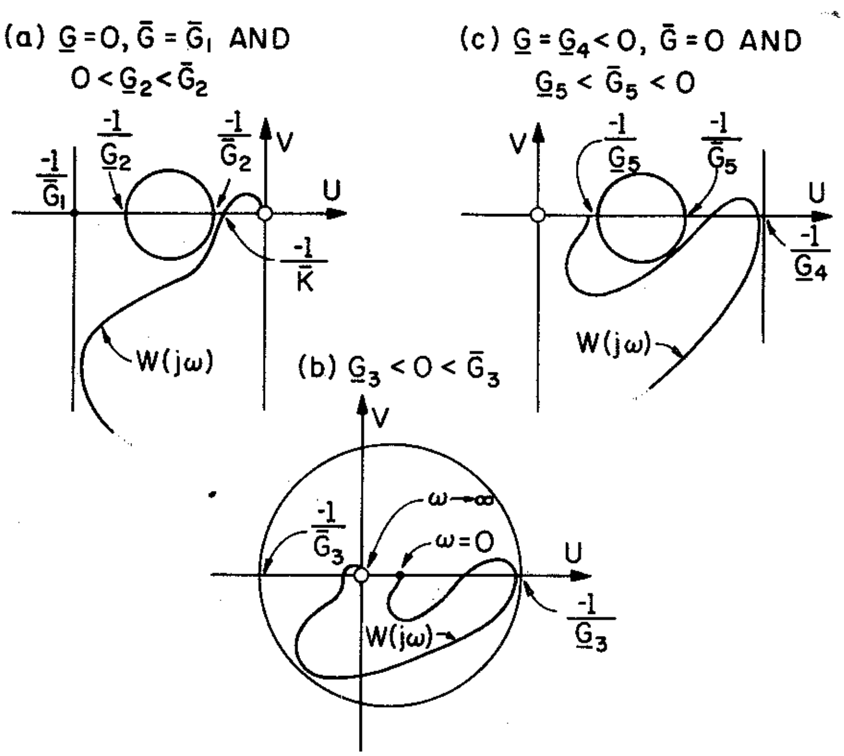


Fig. 3. Illustrations of the Circle Criterion for UASIL (BIBOS)

crossing points, of slope $-\alpha$ and α respectively.

(iii) At $U = -\frac{1}{2}(\underline{F}^{-1} + \overline{F}^{-1})$, $d\hat{V}/dU = 0$ and $V = -\alpha(\overline{F} - \underline{F})/(4\underline{F}\overline{F})$.

(iv) The intersection of the two straight lines of tangency occurs at $(U = -\frac{1}{2}(\underline{F}^{-1} + \overline{F}^{-1})$, $V = -\alpha(\overline{F} - \underline{F})/(2\underline{F}\overline{F})$, which is just twice the value of $\hat{V}$ on the parabola at the same ordinate.

These relations are shown in Fig. 5 in one example of the application of the parabola criterion. For a given choice of $\underline{F}$ and $\overline{F}$, it is usually advantageous

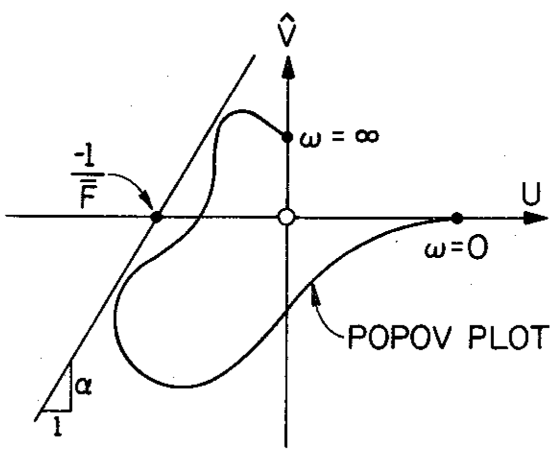


Fig. 4. The Popov Criterion for UASIL (BIBOS)

to choose $|\alpha|$ to be as small as possible, as estimated by drawing the pair of tangency lines. Taking $\alpha < 0$ merely reverses the sense of the parabola.

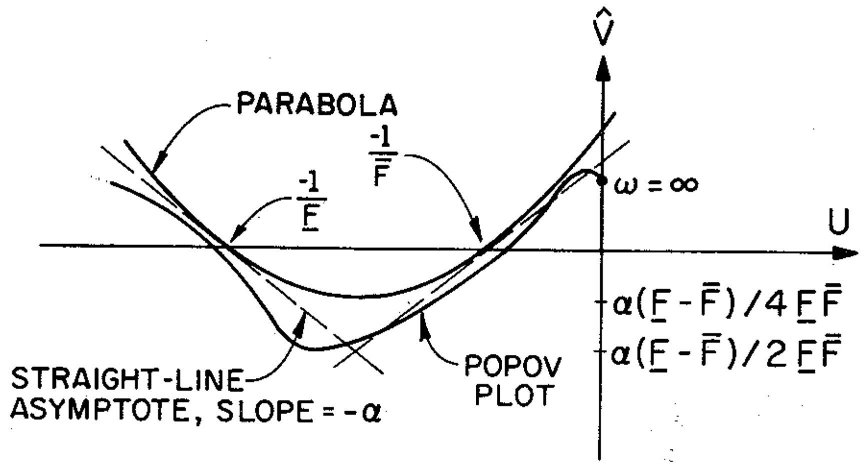


Fig. 5. An Application of the Parabola Criterion

The Off-Axis Circle Criterion (OACC) is also somewhat harder to apply than the CC, because there is one degree of freedom inherent in the choice of the center of the circle. Since the center can usually be located quite quickly by trial-and-error placement of a compass point, there is no need for "rules of thumb" as in the PC above. The special cases $\underline{M} = 0$ or $\bar{M} = 0$ are much simpler, because the OAC degenerates into a straight line that <u>does not need to be vertical</u>. The most important point to be made is that the OAC need only avoid $W(j\omega)$ <u>for positive ω</u>; if this is not kept in mind, then the advantage of the OACC is lost, because the plot of $W(j\omega)$ for $\omega < 0$ is symmetric about the real axis and avoiding that part of the Nyquist plot as well will force the circle center to be on the real axis.

One final caution regarding both circle criteria: It is essential that the vertical and horizontal axes have the same scales; otherwise you are actually applying "ellipse" criteria and, by scaling the V axis appropriately, you could fallaciously "prove" stability for the Hurwitz sector.

As final illustrations of these criteria, consider the following examples, taken directly from [6, Chapter 7].

<u>Example 1</u>: Given a feedback system described by Fig. 1 with the linear plant having a transfer function

$$W(s) = \frac{s + 1}{s(s + 0.1)(s^2 + 0.5s + 9)}$$

(a) The Nyquist plot of $W(i\omega)$ shown in Fig. 6 indicates that the Hurwitz range $(\underline{K}, \bar{K})$ for the stability of an LTI feedback system incorporating this plant is (0, 4.28).

(b) Since the Nyquist plot is asymptotic to a vertical line through the point ($U = -10.06$, $V = 0$) as $\omega \to 0$ [analytically, $U(\omega) \geq U(0) = -815/81$], the system is stable for all nonlinear time-varying gains $g(\sigma,t)$ in the sector $[\varepsilon, 0.0993-\varepsilon]$ by the CC, where ε is arbitrarily small.

(c) The modified Nyquist plot is also indicated in Fig. 6. The straight line that is almost tangential to the modified Nyquist plot at two points intersects the negative real axis at (-2.85, 0), yielding a Popov gain upper bound of $\overline{F}$ = 0.35. The system is thus stable for all nonlinear time invariant functions f(σ) in the sector [ε, 0.35] by the PC.

(d) A straight line nearly tangential to the Nyquist plot at two points intersects the negative real axis at $(-\overline{M}^{-1}, 0)$ where $\overline{M}$ = 3.13. The system is consequently stable for all slope-constrained nonlinearities in the sector [ε, 3.13] by the OACC.

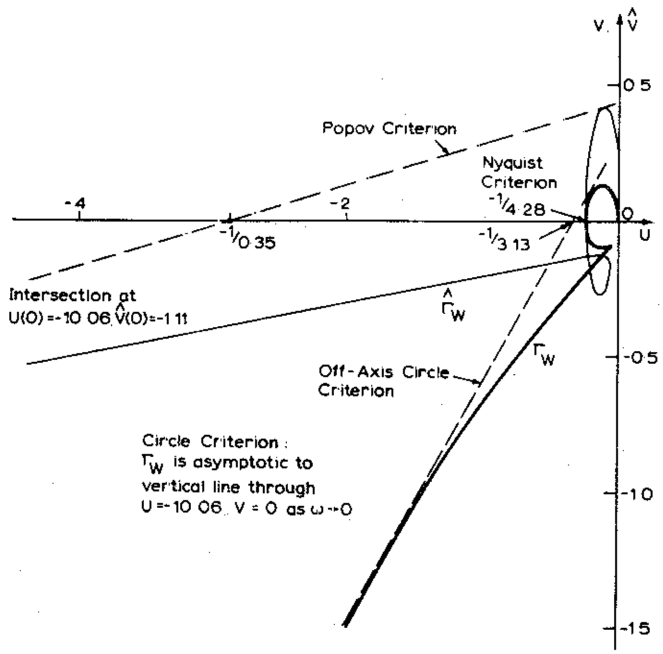


Fig. 6. Example 1: Comparison of Stability Criteria

Example 2: Consider $W(s) = 3(s + 1)/s^2(s^2 + s + 25))$, with the lower bound on the nonlinearity being specified to be unity ($\underline{M} = \underline{F} = \underline{G} = 1$).

(a) The Nyquist criterion applied to the system indicates that the system is stable for all linear gains in the range (0, 8) [Fig. 7].

(b) A circle with its center on the negative real axis is drawn passing through the point (-1, 0) to be nearly tangential to the Nyquist plot. Since it intersects the negative real axis again at $(-2.22^{-1}, 0)$, by the CC the system is stable for all nonlinear and time-varying gains g(σ,t) in the sector [1, 2.22].

(c) A parabola satisfying the PC passing through (-1, 0) and nearly tangential to the Popov plot intersects the negative real axis at U = -0.37. Hence, by the parabola criterion, the system is stable for all nonlinear time invariant gains f(σ) in the sector [1, 2.70].

(d) A circle passing through (-1, 0) and nearly touching the Nyquist curve at two points intersects the negative real axis at $(-7.25^{-1}, 0)$. By the OACC, the system is stable for all slope-constrained gains in the range [1, 7.25].

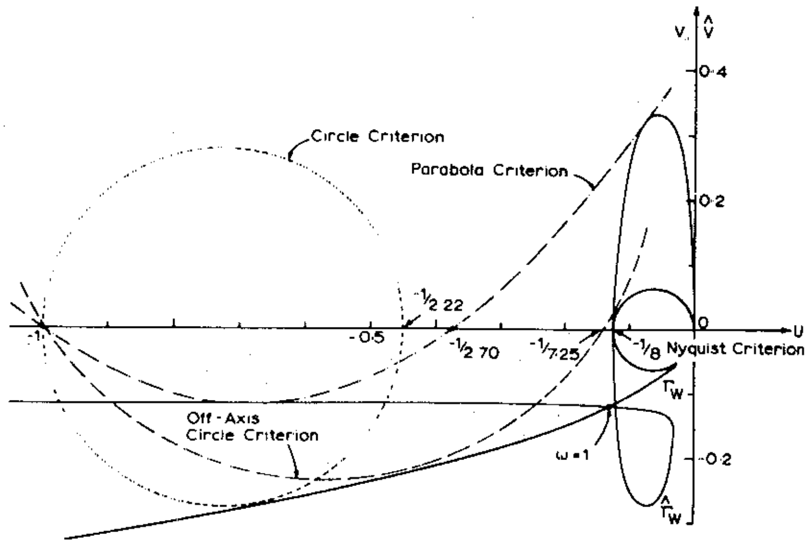


Fig. 7. Example 2: Further comparisons of stability criteria.

To be fair to the CC and PC, it should be observed that Examples 1 and 2 were specifically created to achieve a significant "spread" in the results, i.e., to yield $\overline{G} << \overline{F} << \overline{M}$. To accomplish this, both transfer functions have a pair of very lightly damped complex conjugate poles, which lead to sharp "kinks" in the frequency response plots.

General Comments. It would be very difficult to overemphasize the practical utility of the geometrical stability criteria given in Section 2. The following points bring out the pragmatic elegance of these results.

(a) It is not necessary to know $f(\sigma)$ or $g(\sigma,t)$ exactly. The criteria use only the gross "gain" of the nonlinearity, as defined in Eq. (6), (7), or (9). This is important, because (i) $f(\sigma)$ or $g(\sigma,t)$ is generally not available with any precision [you can't measure these functions for $-\infty < \sigma < \infty$ (and possibly for all time)], (ii) because different elements of the same type (e.g. servo motors) differ from individual to individual, and (iii) because you are not required to make analytic modeling judgments, such as "$f(\sigma) \cong k\sigma^3$" that leave your analysis open to question.

(b) It is not necessary to know $W(s)$ exactly. The criteria only use the basic behavior of $W(j\omega)$, which is important because (i) you could use frequency response test data (e.g. ω_i, $|W(j\omega_i)|$ and $\measuredangle W(j\omega_i)$, $i=1,2,\ldots$) directly without any intervening modeling, and (ii) the results are independent of system order.

(c) They are easy to apply. The application of the CC and OACC are trivial, given the Nyquist plot of $W(j\omega)$ which would be required to determine the stability of the linearized version of the system. The added plot needed for the PC adds very little to the effort.

(d) They are completely rigorous. Unlike small-signal linearization or describing function methods, UASII (BIBOS) is guaranteed.

The formulations of the geometrical criteria given in Section 2 have been made as simple as possible, by leaving out numerous special cases that arise when the circles, lines or parabolas touch the frequency response plots at various points. This results in a very slight "loss of generality" (e.g. in Ex. 1 part (c), the PC could be used to obtain $0 < \frac{f(\sigma)}{\sigma} \le 0.35$ as in [6], instead of $0 < \varepsilon \le \frac{f(\sigma)}{\sigma} \le 0.35$) but no loss of rigor.

4. A REAL-WORLD APPLICATION

The examples outlined in Section 3 are of the "academic type", i.e., specifically designed to illustrate the applications of, and differences between, various stability criteria. In this section, we will consider a case that might occur in an actual control system design problem. The system to be considered is a rotational position control system as depicted in Fig. 8.

The basic plant is modeled simply by a double integration, representing the ideal frictionless equation $\ddot{\theta} = T/I$ where $\ddot{\theta}$ is the angular acceleration of the object being positioned, I is its moment of inertia and T is the applied torque,

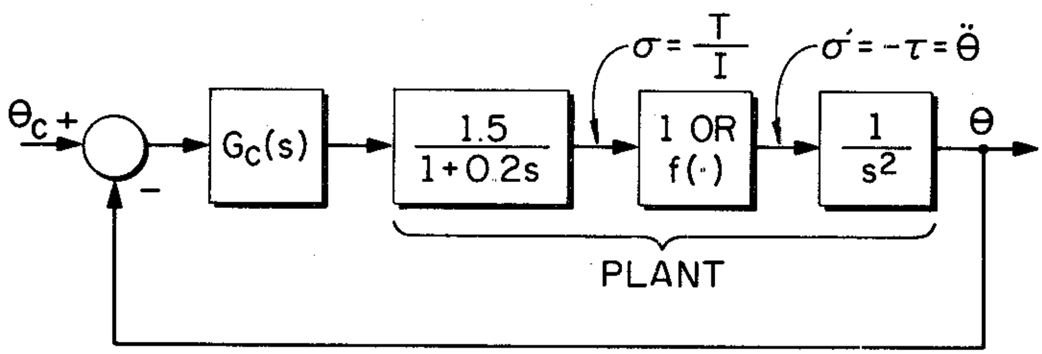


Fig. 8. Rotational Position Control System Model

and a transfer function $1.5/(1 + 0.2s)$, which we assume models the torque amplifier gain divided by I and the dominant lag in the torque source. The output of this block is $\sigma \triangleq T/I$; if the torquer is linear, then the variable σ' in Fig. 8 is $\sigma' = \sigma$, otherwise $\sigma' = f(\sigma)$ or $\sigma' = g(\sigma,t)$. A standard problem in classical linear control theory is to design a compensator, characterized by $G_c(s)$, so that the closed loop system in Fig. 8 with $\sigma' = \sigma$ will meet a given set of design specifications.

The compensator must be of the standard "lead" type, or perhaps of the lead-lag type, because the closed-loop system without compensation is unstable for any positive value of constant gain ($G_c = K$) due to the negative phase margin, as shown in Fig. 9. Following the design specifications $M_m \cong 1.3$, $\omega_m \cong 0.8$ used in D'Azzo and Houpis [7] for this problem (see footnote*, next page), the compensator they obtain leads to a total forward path transfer function of

$$G_t(s) = G_p(s)\ G_c(s) = \frac{0.153\ (1 + 5.55s)}{s^2(1 + 0.2s)(1 + 0.5s)}$$

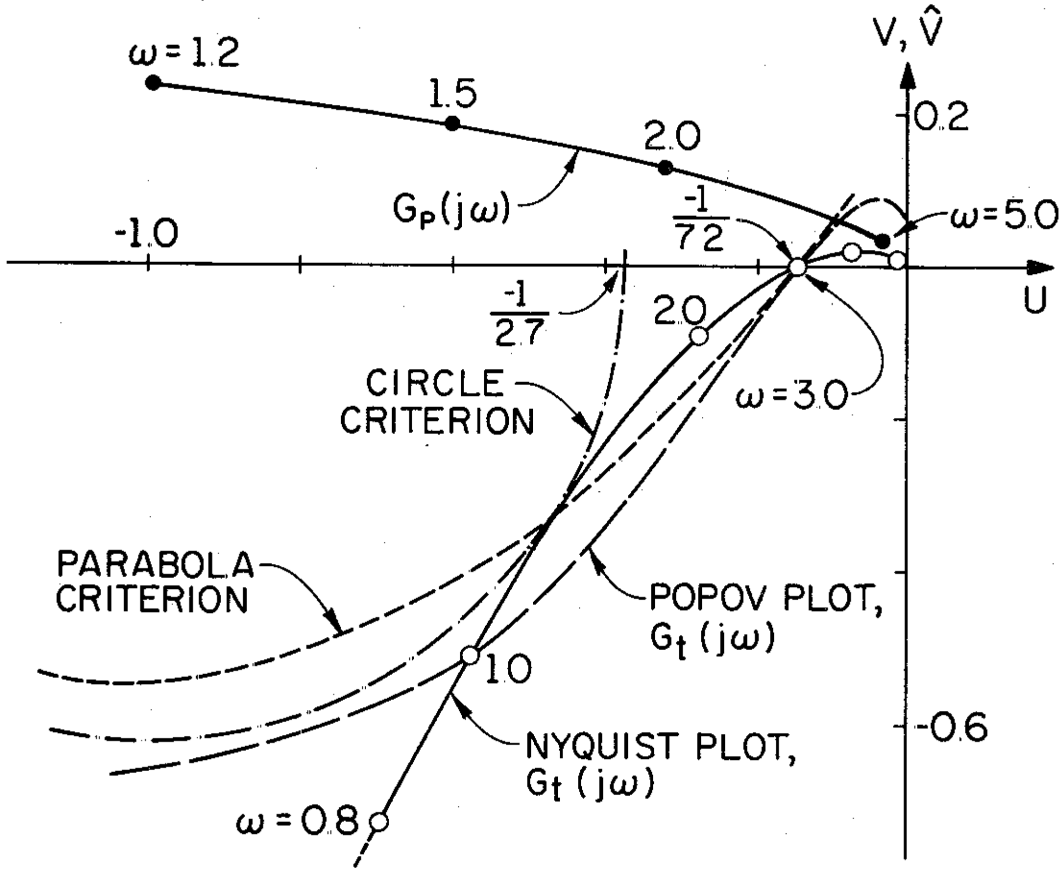


Fig. 9. Nyquist Plot of the Plant $G_p(j\omega)$; Nyquist and Popov Plots of the Compensated Plant $G_t(j\omega)$.

If the torque source is nonlinear and/or time-varying. i.e., if $\sigma' = g(\sigma,t)$ in Fig. 8, it is important to check the design to see how much nonlinearity or time variation can be tolerated. Thus, we construct both the Nyquist plot and Popov plot of $G_t(s)$, as shown in Fig. 9. Before considering the nonlinear case, note that the Hurwitz range (or maximum value by which the gain can be increased) is $\overline{K} = 7.2$, as the Nyquist plot shows. As a first investigation, let us seek $\underline{G}$, $\overline{G}$ by the circle criterion and $\underline{F}$, $\overline{F}$ by the parabola criterion such that $-1/\underline{G}$ and $-1/\overline{G}$ are symmetric about 1, and similarly for $\underline{F}$, $\overline{F}$ -- i.e., let us find the largest circle centered on U = -1 so that the Nyquist plot of $G_t(j\omega)$ is not intersected, and the largest parabola with U = -1 as its axis so that the Popov plot of $G_t(j\omega)$ is not touched. From Fig. 9 we obtain the following bounds:

CC: $\underline{G} = 0.61$, $\overline{G} = 2.70$

PC: $\underline{F} = 0.537$, $\overline{F} = 7.20-\varepsilon$ ($\cong \overline{K}$)

* These parameters are the classical "M-peak" or maximum value of $M = |G_cG_p(j\omega)|/|1 + G_cG_p(j\omega)|$, and the frequency at which M_m occurs, which is a measure of the closed-loop system bandwidth.

These constraints are depicted in Fig. 10, for several hypothetical symmetric nonlinearities. If $g_1(\sigma,t_1)$ is typical of the nonlinear characteristic, then there is no problem with respect to stability, even if $g_1(\sigma,t)$ is rapidly time-varying as long as it always lies in the sector for all σ,t. If, however, $f_2(\sigma)$ is typical, then time variation is potentially a problem (i.e., we have no guarantee of stability) while we are assured of stability if f_2 is time invariant.

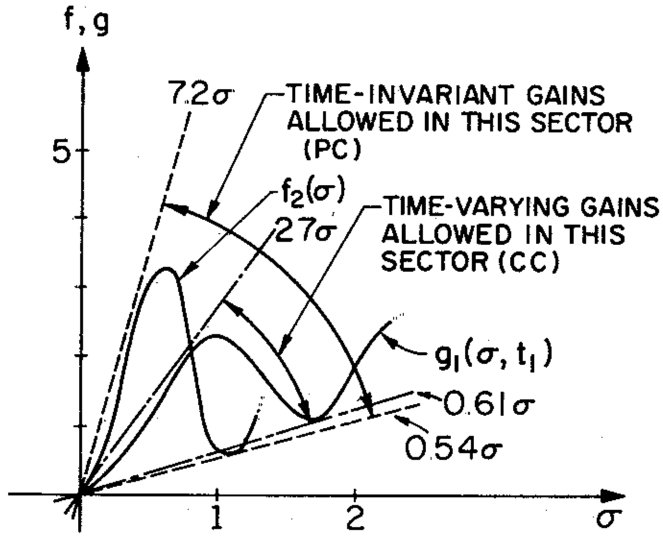


Fig. 10. Gain Sector Constraints from the CC and PC.

The above results (Fig. 10) are not unique -- they were obtained on the basis of the general desire to find stability regions that are geometrically centered about the ideal characteristic $\sigma' = \sigma$. Suppose the torque source is basically saturating, in the sense that $g(\sigma,t)/\sigma$ or $f(\sigma)/\sigma \leq 1$? Then we have the condition $\overline{F} = \overline{G} = 1.0$ and we seek the minimum $\underline{G}$ and $\underline{F}$ permitted by the CC and PC, respectively.

Determining $\underline{G}$ is very simple, as shown in Fig. 11. A circle is drawn through $-1/\overline{G} = -1$ that just avoids the Nyquist plot of $G_t(j\omega)$; the other intersection with the real axis is at $-1/\underline{G}$, so

$$-1/\underline{G} = -7 \rightarrow \underline{G} = \frac{1}{7.0} = 0.14$$

To find the lower bound $\underline{F}$ by the PC, we first observe in Fig. 11 that $\hat{V}$ asymptotically approaches -0.74 as ω approaches 0. We then draw a straight line through $-1/\overline{F} = -1$ which just misses the Popov plot for $\omega \cong 5$. This straight line passes through $\hat{V} = -0.74$ at $U = -6.6$, as shown, so by extrapolation it will reach $\hat{V} = -1.48$ at $U = -12.2$. The parabola for which this line is one asymptote (see the discussion on applying the PC in Section 3) thus can have its minimum at $U_0 = -12.2$, $\hat{V}_0 = -0.74$ as required; furthermore, by symmetry, this parabola will intersect the negative real axis again at $-1-2(11.2) = -23.4$ which gives us $\underline{F}$:

$$-1/\underline{F} = -23.4 \rightarrow \underline{F} = \frac{1}{23.4} = 0.043$$

To interpret these results, the saturation curve $g_1(\sigma,t_1)$ shown in Fig. 12 is no problem even if there is rapid time variation within the sector $(\underline{G}, \overline{G})$, while $f_2(\sigma)$ is acceptable only if the torque characteristic is time invariant.

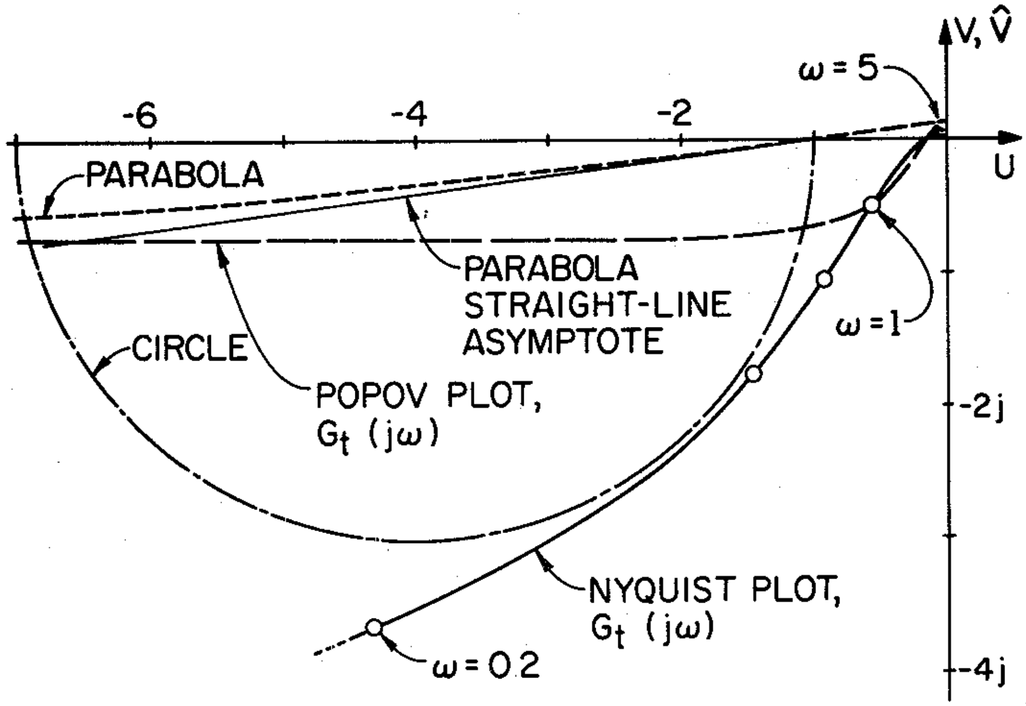


Fig. 11. Obtaining Gain Sectors for a Saturating Torquer by the PC and CC.

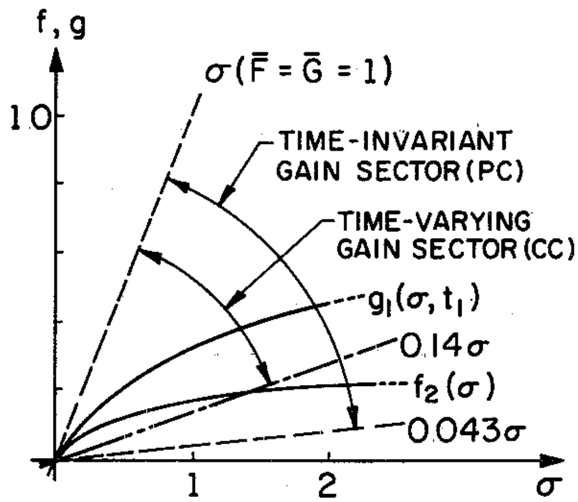


Fig. 12. Nonlinearity Constraints for Saturating Iorquers.

The stability regions or sectors depicted in Figs. 10 and 12 show that there is a substantial degree of nonlinearity that can be allowed with a complete assurance of UASIL or BIBOS. In part, this is due to the conservatism of the classical control theory design procedures; the specifications $M_m = 1.3$ led to a gain margin of $\bar{K} = 7.2$ (Fig. 9) which from the linear system standpoint is indeed prudent. The fact that the PC can give us an absolute guarantee of stability for $0.54 \leq \frac{f(\sigma)}{\sigma} \leq 7.2-\epsilon$ where ϵ is arbitrarily small is, in the author's opinion, quite remarkable. Furthermore, if the nonlinear torquer is monotonic (slope constrained), then the OACC guarantees UASIl for the entire Hurwitz range, $0 \leq df/d\sigma \leq 7.2-\epsilon$. On the whole, the above example should provide considerable insight into the use and power of the basic stability criteria of Section 2.

5. USE OF ABSOLUTE STABILITY CRITERIA IN NONLINEAR SYSTEMS DESIGN

In the problem considered in Section 4, it is assumed that the system design was completed using the linear control theory approach, and then the possible destabilizing effect of a nonlinear torquer characteristic was explored. A more sophisticated approach would be to use absolute stability criteria during the design phase. The basic ideas can be conveyed by a simple example; engineers versed in the basics of classical control theory can readily extend the methodology to other, possibly more complicated, situations.

The closed-loop system to be designed is again the rotational position control system shown in Fig. 8, with one difference: the dominant nonlinear effect is a gain-changing nonlinearity, used to null large angular position errors rapidly without excessive overshoot (Fig. 13); for small errors, $|\sigma| \leq \delta$, the output is σ, while for larger errors, the slope of the nonlinearity is K, as shown. The system design will then involve choosing $G_c'(s)$, an appropriate compensator, so that

(a) The control system meets the usual system specifications (e.g. $M_m = 1.3$) for small errors (i.e. with $f(\sigma) = \sigma$)

(b) The control system is UASIL (BIBOS) when the gain changing nonlinearity is taken into account.

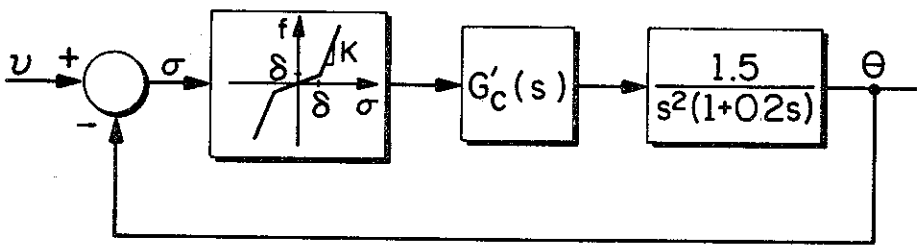


Fig. 13. Rotational Position Control System with Gain-Changing Nonlinearity

The simplest design problem is obtained when the system is assumed to be nonlinear and time-varying, which necessitates using the CC. (The procedure is most direct, because there is no freedom in drawing the circle through $-1/\underline{G}$ and $-1/\bar{G}$.) Then, determining G_c' simply involves forcing $G_t' = G_c'G_p$ to avoid two circles: The $M_m = 1.3$ circle (cf. D'Azzo and Houpis [7] Chapter 11) and the CC circle, passing through -1 and -1/K for the nonlinearity shown in Fig. 13. Taking K to be 5, we see that the compensator design G_c of Section 4 will fail, in the sense that $G_pG_c(j\omega)$ cuts the CC circle, as shown in Fig. 14. However, modifying the compensator or adding a second one so that the total linear part frequency response plot avoids the CC circle, as illustrated in Fig. 14, is a straightforward application of the classical frequency domain cascade compensator methods.

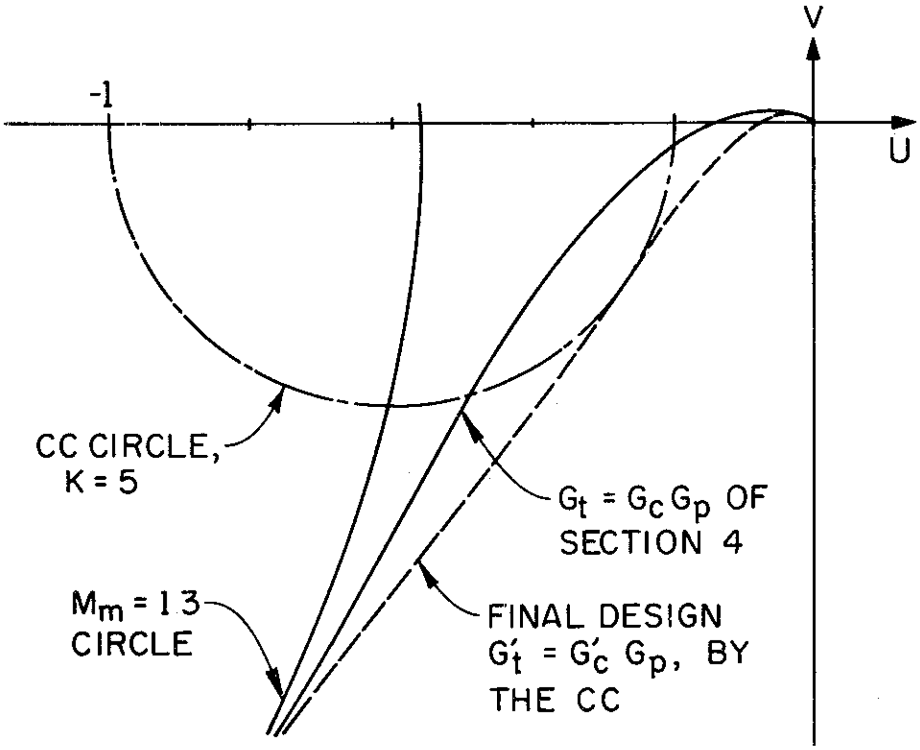


Fig. 14. System Design Using the CC

Of course, the above design procedure (avoiding both the CC and M_m circles by suitable compensator choice) will be valid if the system is time invariant as well. However, a design that is often much less conservative will be obtained by using the PC or OACC. For example, we saw in Section 4 that the compensated system is stable for K as large as 7.2 as long as the system is not time-varying. If K is greater than 7.2, then a suitable design can usually be found quickly by drawing an off-axis circle through -1 and -1/K which matches the slope of the first-cut compensated Nyquist plot (Fig. 9) near the real axis crossing, and design a new compensator G_c' so that the Nyquist plot of $G_cG_p'(j\omega)$ avoids this OAC. This approach is outlined in Fig. 15, where the final $G_cG_p'(j\omega)$ curve is sketched to satisfy the above constraints, for K = 10.

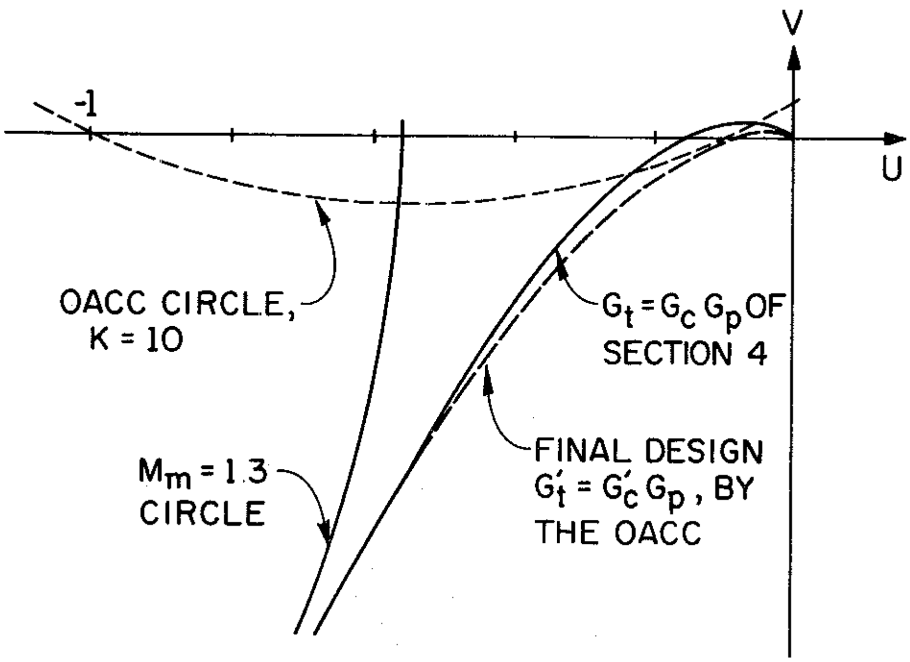


Fig. 15. System Design Using the OACC

The design procedures outlined in this section are fairly informal, and rely heavily on the older frequency domain approach to control systems synthesis. (Which, it might be added, is enjoying a modest revival, as the modern control community comes to appreciate the value of such methods.) The main intent of the examples sketched in Figs. 14 and 15 is to illustrate how control systems can be designed safely and quite simply without neglecting dominant, possibly destabilizing, nonlinear and/or time-varying effects.

6. MORE SOPHISTICATED CRITERIA FOR TIME-VARYING SYSTEMS

The main thrust of this presentation is the rigorous analysis of nonlinear systems. To this point, time variation has only been considered gratuitously, as an added benefit of applying the circle criterion. In effect, the criteria of Sections 2–5 represent an all-or-nothing situation, i.e., gains may either vary as rapidly as you like (the CC), or not vary at all (the PC, OACC). As mentioned in [1], systems in which the nonlinear and/or time-varying part is described by

(i) $\tau = -k(t)\sigma$

or

(ii) $\tau = -k(t)f(\sigma)$

can often be treated more effectively by using an extension of the PC. In essence, this extension allows the analyst to state precisely how rapidly k(t) can vary while guaranteeing UASIL (BIBOS).

In the linear case, (i), the stability conditions are most simply stated for a lower zero bound,

$$0 \le k(t) \le \overline{K}$$

As always, the corresponding linear time-invariant system must be asymptotically stable for $0 \le k \le \overline{K}$; then the stability conditions are that the Popov condition is satisfied by $W(j\omega)$, and $\frac{dk}{dt}$ is restricted.

<u>Theorem 5</u>: A system of the form portrayed in Fig. 1 is UASIL (BIBOS) if $\tau = -k(t)\sigma$, $0 \le k(t) \le \overline{K}$ is within a Hurwitz range, and

(a) The Popov plot of $W(j\omega)$ does not touch or intersect a straight line through $U = -1/\overline{K}$, $\hat{V} = 0$ with slope α, where α is any real non-zero number;

(b) $\frac{dk}{dt} \le 2\alpha\, k(t)[1 - k(t)/\overline{K}]$ if $\alpha > 0$,

or (11)

$\frac{dk}{dt} \ge 2\alpha\, k(t)[1 - k(t)/\overline{K}]$ if $\alpha < 0$ ●

If a "phase-plane portrait" or plot of dk/dt versus k is drawn, then the above constraint on k can be visualized very simply, as shown in Fig. 16.

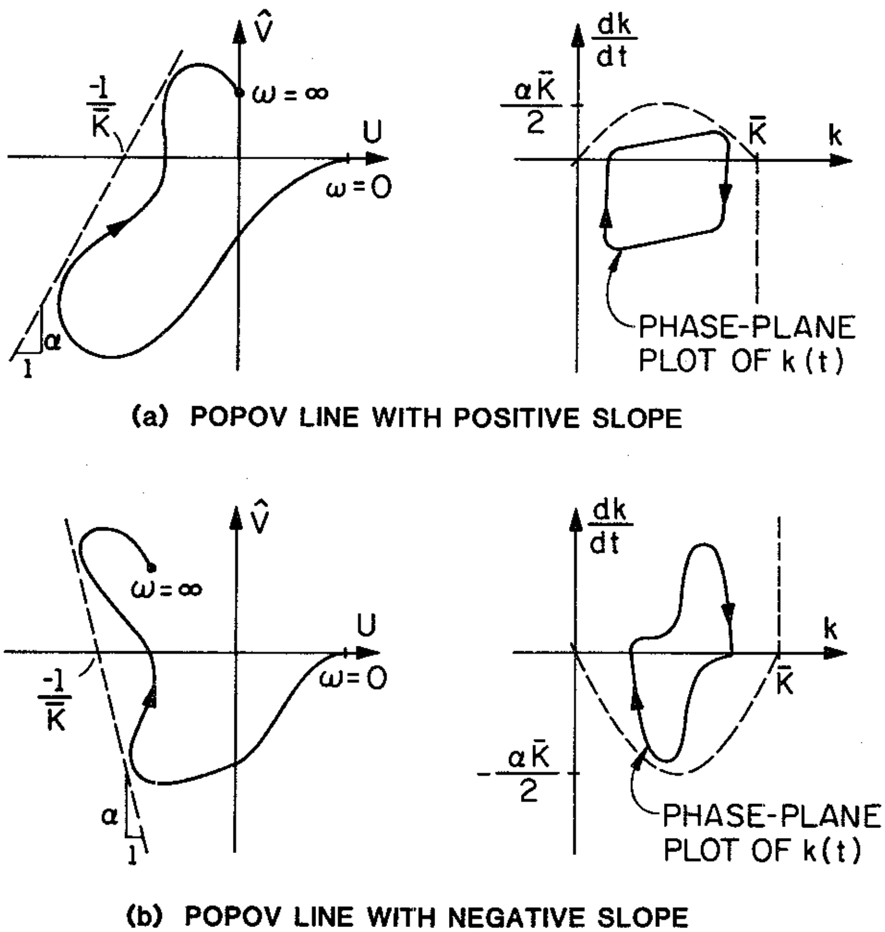


Fig. 16. Geometrical Interpretation of Theorem 5

Comments

(1) It is rare that $k(t)$ actually occupies the entire range $[0,\bar{K}]$. In general, the shape of the phase plane portrait is not compatible with the parabolic constraint, so usually $K_1 \leq k(t) \leq K_2$ where $0 < K_1 < K_2 < \bar{K}$. This is particularly true when dealing with periodic functions: The phase plane trajectory of a periodic function must cross the k-axis with infinite slope, which is not permitted by the parabolic restriction on dk/dt at 0 and $\bar{K}$. Thus in both examples indicated in Fig. 16 $k(t)$ occupies a range that is smaller than $[0,\bar{K}]$.

(2) The constraint on dk/dt can be applied in several ways. The shape of the phase plane plot of k is determined by the wave form in the time domain; for example, sinusoidal gains correspond to ellipses in the phase plane, while nonlinear oscillations are irregular (see Fig. 16). The horizontal width of the trajectory corresponds to the range $[K_1,K_2]$ of $k(t)$, while the vertical dimension is proportional to the range and frequency. Thus if the range is given, we extend the trajectory only in the vertical direction until it touches the parabola to determine the maximum allowed frequency. If the frequency is fixed, then the portrait is enlarged until it touches the parabola which determines the maximum range permitted by the constraint.

(3) If the Popov line is vertical ($\alpha = \infty$), then constraint (11) is removed; the gain can vary between 0 and $\bar{K}$ arbitrarily rapidly. The circle criterion applied to linear time-varying systems is thus a special case of Theorem 5.

(4) The generalization of Theorem 5 to case (ii) above, $\tau = -k(t)f(\sigma)$, simply involves replacing the factor 2 in Eq. (11) with ϕ defined by

$$\phi \triangleq \min_{\sigma} \left\{ \frac{\sigma f(\sigma)}{\int_0^{\sigma} f(\xi)\, d\xi} \right\}$$

the factor ϕ is in some sense a "measure" of the nonlinearity of $f(\sigma)$; if $f(\sigma) = \sigma$ then $\phi = 2$, if $\frac{df}{d\sigma} \geq 0$ then $\phi \geq 1$, and if $f(\sigma)$ is a power law relation we have

$$f(\sigma) = k\, |\sigma|^{\beta} \operatorname{sign}(\sigma) \quad \rightarrow \quad \phi = \beta + 1$$

where $\beta \geq 1$ (see Ref. [6], p. 30). It is interesting to note that a cubic nonlinearity, $f(\sigma) = \sigma^3$, would permit $k(t)$ to vary twice as rapidly since $\phi = 4$, than in the linear case (ϕ=2)!

The generalizations of the Popov criterion given in Theorem 5 and point (4) above constrain dk/dt at every instant. There is another, generally substantially less strict, time-averaged constraint (integral constraint) on dk/dt that was obtained using a generalization of the Lyapunov direct method (Taylor and Narendra [8]) which is discussed in detail in Ref. [6].

The applications of Theorems 2 and 5 and of the more powerful result mentioned above to simple second-order linear and nonlinear time-varying system equations, viz.

$$\ddot{y} + 2\zeta\dot{y} + [a - 2q \cos(2t)]y = 0 \qquad (12)$$

(the damped Mathieu equation) or

$$\ddot{y} + 2\zeta\dot{y} + \delta^2 y + [a - \delta^2 - 2q \cos (2t)]\, f(y) = 0$$

(a nonlinear form of Mathieu's equation) is considered in detail in Chapter VIII of Ref. [6].

Finally, it should be pointed out that the application of stability criteria for linear or nonlinear time-varying systems is even more important than in the nonlinear time-invariant case. This is due to the fact that one can often "get away with" using small signal linearization, describing function methods or Aizerman's conjecture for nonlinear time invariant systems, but for time-varying systems the assumption that "the system is UASIL for slowly-varying gains" is fuzzier and more dangerous. The following examples illustrate this point vividly.

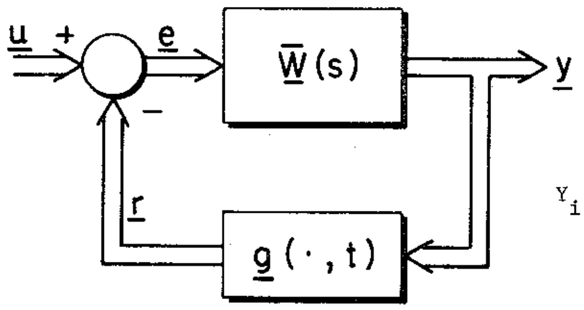
(i) $\dot{y} + k(y)y = 0$ <u>is</u> UASIL for $k(y) > 0$, while
$\dot{y} + k(t)y = 0$ may not be UASIL for $k(t) > 0$; e.g.,
if $k(t) = e^{-t}$ then $y = y_o \exp (e^{-t} - 1)$ which approaches y_o/e.

(ii) $\ddot{y} + 2\zeta\dot{y} + k(y)y = 0$ <u>is</u> UASIL for $k(y) > 0$, while for small ζ Eq. (12) with $a=1$ is UASIL only if $|q| < 2\zeta$ ([6], p. 187) which is more restrictive than $|q| < \frac{1}{2}$ which makes $k(t) > 0$.

Thus, although Theorem 5 and related results (see Ref. [6]) may be somewhat more difficult to apply than Theorems 3 and 4, it may be appropriate to place greater emphasis on the time-varying case. (Of course, the circle criterion is a very powerful result for time-varying systems; it is often too strict, however, because dk/dt is unconstrained. For a comparison of Theorems 2 and 5, see [6].)

7. MULTIPLE NONLINEARITIES

An obvious area of extension is the case where more than one nonlinear or time-varying effect is important. A great deal of generality can be achieved by considering the basic system model of Fig. 1 to be <u>vectorial</u> - i.e., each "signal" in the closed loop system is a vector made up of m variables, as shown in Fig. 17. Thus the linear part of the system is represented by an m x m matrix of transfer functions, $\bar{W}(s) = [W_{ij}(s)]$, and it is assumed that the nonlinear time-varying gains are "decoupled," so each variable r_i is a function only of y_i and t.



$$Y_i(s) = \sum_{j=1}^{m} W_{ij}(s)\, [U_i(s) - R_i(s)]$$

$$r_i = g_i\,(r_i,\ t),\ i = 1,2,\ldots,m$$

Fig. 17. A Generalized Model for Multiple Nonlinearities

Some stability results that exist for such systems are summarized in Ch. 9 of [6]. It is not surprising that the concepts generalize quite directly, e.g., there is a direct analytic analog to the CC, but the applications become much more difficult. There are a few special cases where the stability conditions can be formulated as geometric constraints on the elements of $\overline{W}(j\omega)$; in general, however, the conditions must be dealt with analytically, which can be very tedious. For this reason, we restrict our attention to one example that is probably the most elegant application of this branch of stability theory. The details of this problem are given in Ch. 9 of [6] and in [9].

In a classical problem of optimal control, it can be shown that an integrated quadratic performance index is minimized by using linear state-variable feedback, i.e., by feeding back a signal such as $\underline{y}$ in Fig. 17 which is a linear combination of the states:

$$\underline{e} = \underline{u} - \underline{y}, \qquad \underline{y} = \overline{M}\,\underline{x}$$

where $\overline{M}$ is an (m x n) matrix of constant gains (generally m=n), and $\underline{x}$ is a state vector made up of the variables that characterize the plant dynamic behavior, as in Eq. (2). In other words, the solution to the optimal control problem can be represented in the form shown in Fig. 17 with a unity gain in each feedback path, $r_i = y_i$. The gain matrix $\overline{M}$ is determined by the quadratic performance index (see Ref. [6]) and governs $\overline{W}(s)$; the exact solution is not needed for this discussion.

The stability result is the following: If $\overline{W}(s)$ in Fig. 17 represents the solution to the optimal control problem outlined above, then the nonlinear closed-loop system with each nonlinear time-varying characteristic satisfying

$$\frac{1}{2} \leq \frac{g(y_i,t)}{y_i} < \infty$$

in UASIL. This surprising result demonstrates that the optimal control system is very robust, in the sense that the state-variable feedback can differ substantially from the ideal or exact solution $g(y_i,t) = y_i$ without a loss of stability. This result was first given in Ref. [6]; in the time-invariant case, the condition above was obtained by Moore and Anderson [9].

8. SUMMARY AND CONCLUSIONS

The main objective of this presentation is to focus on the simplest rigorous stability criteria that exist, to illustrate and motivate their use. Since the frequency domain representation of linear systems dynamics is quite prevalent, especially among the practicing control engineering community, it should be readily appreciated that stability analyses such as those outlined in Sections 3 and 4 can be performed quite simply -- certainly, with little added effort if Nyquist's criterion has already been applied to a linearized version of the system.

The major philosophical reservation that seems to have discouraged the wide-spread use of the criteria given in Section 2 -- that the stability conditions obtained are often conservative, as mentioned in the introduction -- seems to result from an attempt to obtain more from the stability criteria than can reasonably be expected, or needed. Take the problem treated in Section 4 as a case in point.

The control system in Fig. 8 is designed so that the unity-gain linearized model, $g(\sigma,t) = \sigma$, satisfies certain performance specifications. Various stability criteria were then applied to obtain the stability results summarized in Table 1. First, observe that the parabola criterion (PC) result is not terribly conservative; the nonlinearity gain (in the sense shown in Table 1) may be as close to the linear gain limit ($\overline{K} = 7.2$) as desired, and can be substantially less than unity ($\underline{F} = 0.54$). The circle criterion (CC) result, on the other hand, does appear to be conservative in the sense that g/σ is not allowed to exceed 2.7, which is much less than $\overline{K}$. Realistically, however, even if the system were linear time-invariant, its performance will be degraded considerably if k exceeds 2.7; in fact, the M-peak value for k = 2.7 is $M_m = 3.5$ which fails to

TABLE 1 – Summary of Stability Results from Section 4

Feedback Law	Gain Constraints	Comments
$\sigma' = k\sigma$	k=1	Linear time-invariant; meets performance specifications
$\sigma' = k\sigma$	$0 \leq k < 7.2$	Linear time-invariant; asymptotically stable
$\sigma' = f(\sigma)$	$0.54 \leq \frac{f(\sigma)}{\sigma} \leq 7.2-\varepsilon$	Nonlinear time-invariant; PC→UASIL (ε may be arbitrarily small)
$\sigma' = g(\sigma,t)$	$0.61 \leq \frac{g(\sigma,t)}{\sigma} \leq 2.7$	Nonlinear time-varying; CC→UASIL
$\sigma' = f(\sigma)$	$0 \leq \frac{df}{d\sigma} \leq 7.2-\varepsilon$	Monotonic nonlinear time-invariant; OACC→UASIL

meet the design specification $M_m = 1.3$ by a considerable margin. Thus, since the linear system with k > 2.7 does not achieve the desired performance, it is doubtful that a nonlinear time-varying system with $\overline{G} \geq 2.7$ would be desirable even if it were "safe" from the stability standpoint. In this sense, the CC result may very well be all that is required. Finally, imposing the much more stringent slope restriction on the nonlinearity $0 \leq \frac{df}{d\sigma} \leq 7.2-\varepsilon$ by the OACC leads to a stability condition that is not at all conservative.

It is the author's hope that this presentation will encourage greater interest in the use of absolute stability criteria by practicing engineers, and that the applications literature in this area will be enriched thereby.

REFERENCES

[1] Hedrick, J. K. and Paynter, H. M. (Eds.), Nonlinear System Analysis and Synthesis: Vol. 1 – Fundamental Principles, Chapter 2, American Society of Mechanical Engineers, 1978.

[2] Kalman, R. E. and Bertram, J. E., "Control System Design Via the Second Method of Lyapunov," J. Basic Engrg., Trans. ASME, Vol. 82, No. 2, 1960.

[3] LaSalle, J. P. and Lefschetz, S., Stability by Lyapunov's Direct Method with Applications, Academic Press, New York, 1961.

[4] Holtzman, J. M., Nonlinear System Theory, Prentice-Hall, Englewood Cliffs, New Jersey, 1970.

[5] Vidyasagar, M., Nonlinear Systems Analysis, Prentice-Hall, Englewood Cliffs, New Jersey, 1978.

[6] Narendra, K. S. and Taylor, J. H., Frequency Domain Criteria for Absolute Stability, Academic Press, New York, 1973.

[7] D'Azzo, J. J. and Houpis, C. H., Feedback Control System Analysis and Synthesis, McGraw-Hill, New York, 1966 (Second Ed.).

[8] Taylor, J. H. and Narendra, K. S. "The Corduneanu-Popov Approach to the Stability of Nonlinear Time-Varying Systems," SIAM J. on Appl. Math, Vol. 18, No. 2, 1970.

[9] Moore, J. B. and Anderson, B. D. O., "Applications of the Multivariable Popov Criterion," Int. J. of Control, Vol. 5, No. 4, 1967.