STABILITY THEORY: INCLUDING SATURATION EFFECTS

The basic concept of the stability of dynamical systems is carefully developed, and its application to the design of feedback control systems is presented. Classical stability theories for linear and nonlinear feedback control systems are discussed and the Lyapunov and Routh–Hurwitz stability criteria are developed. Next, attention is given to conventional control design and to frequency-domain methods: The root locus and Nyquist stability criteria are presented, and the Bode plot and Nichols chart-based methods for stability and for degree of stability determination are discussed. The emphasis then shifts to the stability of control systems comprising linear plants and actuation elements which are subject to saturation. Then, nonlinear controllers are called for. Moreover, the broader question of bounded input and bounded output stability is addressed. Thus, a comprehensive time-domain

J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*.

design methodology for nonlinear tracking controllers for actuator saturation effects mitigation, is presented.

LYAPUNOV STABILITY CRITERION

Stability is a very important characteristic of a system. If a system is linear and time-invariant, several criteria such as the Nyquist and Routhian stability criteria may be applied to determine stability. However, for nonlinear and time-varying systems, the Lyapunov direct method is applicable. Its advantage is that it is not necessary to solve the system equation, which is generally very difficult.

The system with zero input is defined by

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \tag{1}$$

and contains the n-dimensional response vector $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$. The vector $\mathbf{f}(\mathbf{x}, t)$ is a function of $\mathbf{x}$ and t. The solution of Eq. (1) is a function of the initial condition $\mathbf{x}(0)$ and the initial time t_0 and is given by

$$\mathbf{x} = \phi(t, \mathbf{x}(0), t_0) \tag{2}$$

The length of the vector $\mathbf{x}$ can be measured by the Euclidean norm (1)

$$\|\mathbf{x}\| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2} \tag{3}$$

The system of Eq. (1) is uniformly stable (2) when there is a finite positive constant ϵ such that, for any $\mathbf{x}(0)$ and t_0, the solution of Eq. (2) satisfies the condition

$$\|\mathbf{x}\| \leq \epsilon \|\mathbf{x}_0\|, \qquad t \geq t_0 \tag{4}$$

where $\mathbf{x}_0$ represents the equilibrium solution as $t \to \infty$. The vector $\mathbf{x}_0$ may be a point or it may be a continuous oscillation (limit cycle). Figure 1 represents the graphical plots of several possible solutions of Eq. (2) for an initial condition $\mathbf{x}(0)$ and a finite final value $\mathbf{x}_0$.

Stability in the Sense of Lyapunov

Consider a region ϵ in the state space enclosing an equilibrium point $\mathbf{x}_0$. This equilibrium point is stable provided that there is a region $\delta(\epsilon)$, which is contained within ϵ such that any trajectory (3) starting in the region δ does not leave the region ϵ. Note that with this definition it is not necessary for the trajectory to approach the equilibrium point. It is necessary only for the trajectory to stay within the region ϵ. This permits the existence of a continuous oscillation about the equilibrium point. The state-space trajectory for such an oscillation is a closed path called a *limit cycle*. The performance specifications for a control system must be used to determine whether or not a limit cycle can be permitted. The amplitude and frequency of the oscillation may influence whether it represents acceptable performance. When limit cycles are not acceptable, more stringent restraints must be chosen to exclude their possible existence (25).

Limit cycles are usually characterized as stable or unstable. If a limit cycle is stable, it means that trajectories in the state space on either side will approach the limit cycle. An unstable limit cycle is one in which the trajectories on either side diverge from the limit cycle. These trajectories may approach other limit cycles or equilibrium points.

In the case that there is an input $\mathbf{u}(t)$, the system is defined by

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \tag{5}$$

In linear systems, stability is a system characteristic which is independent of any initial condition $\mathbf{x}(0)$ and/or the input magnitude $\mathbf{u}(t)$. However, in nonlinear systems the stability of the system may depend on the initial condition and/or the magnitude of the input (4). This case is typical for control systems and is considered later in this article.

An important case is a system described by bounded-input–bounded-output (BIBO) stability. In such a system, either linear or nonlinear, the output response to any bounded input is also bounded. For a linear time-invariant (LTI) system, the necessary and sufficient condition for stability is that (5,6) the integral of the system weighting function or impulse response must be finite. This is expressed by

$$\int |h(t)|\, dt < \infty \tag{6}$$

An equivalent condition for an LTI system is that all eigenvalues (3) are located in the left-half plane. In that case, the transient response associated with those eigenvalues decrease with time, and the output response therefore approaches the particular solution which is determined by the input.

Asymptotic Stability

An equilibrium point is *asymptotically stable* if, in addition to being stable in the sense of Lyapunov, all trajectories approach the equilibrium point. This means that the perturbation solution $\mathbf{x}^*(t)$ approaches $\mathbf{0}$ as time t approaches infinity. This is the stability definition usually used in control-system design.

Figure 1 shows trajectories in the state plane illustrating the general principle of Lyapunov stability (3), asymptotic stability, and instability. The trajectory a starting at the initial state point $\mathbf{x}(0)$ and remaining within the region ϵ meets the conditions for Lyapunov stability. This trajectory is closed, indicating a continuous output oscillation or limit cycle. Trajectory b terminates at the equilibrium point $\mathbf{x}_0$; thus

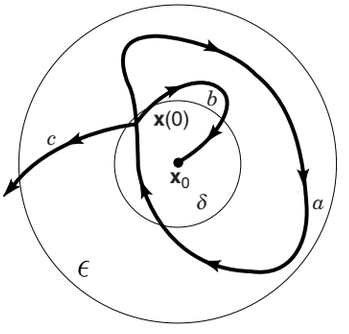

Figure 1. State-plane trajectories indicating (a) Lyapunov stability; (b) asymptotic stability; and (c) instability.

it represents asymptotic stability. Trajectory c leaves the region ϵ and therefore indicates instability.

When the region δ includes the entire state space, the definitions of Lyapunov and asymptotic stability are said to apply in a global sense. The stability or instability of a linear system is global because any initial state yields the same stability determination. A stable linear system is globally asymptotically stable.

The technique for linerarizing nonlinear differential equations in the neighborhood of their singularities or equilibrium points is presented in the section entitled "Linearization (Jacobian Matrix)." The validity of determining stability of the unperturbed solution near the singular points from the linearized equations was developed independently by Poincaré and Lyapunov in 1892. Lyapunov (4,7) designated this as the *first method.* This stability determination is applicable only in a small region near the singularity and results in stability *in the small.* The section on linearization considers Lyapunov's *second method,* which is used to determine stability *in the large*. This larger region may include a finite portion, or sometimes the whole region, of the state space.

QUADRATIC FORMS

Some of the techniques used in determining stability of control systems and for optimizing their response utilize scalar functions expressed in quadratic form. The necessary background for expressing functions in quadratic form is developed in this section. Then some important properties of quadratic forms are presented.

Conjugate Matrix. The elements of a matrix may be complex quantities. For example, a matrix $\mathbf{A}$ may have the elements $a_{ij} = \alpha_{ij} + j\beta_{ij}$. A conjugate matrix $\mathbf{B}$ has elements with the same real component and with imaginary components of the opposite sign; that is, $b_{ij} = \alpha_{ij} - j\beta_{ij}$. This conjugate property is expressed by

$$\mathbf{B} = \mathbf{A}^* \tag{7}$$

Inner Product. The inner product is also called a *scalar* (or *dot*) product since it yields a scalar function. The scalar product of vectors $\mathbf{x}$ and $\mathbf{y}$ is defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle = (\mathbf{x}^*)^{\mathrm{T}}\mathbf{y} = \mathbf{y}^{\mathrm{T}}\mathbf{x} = x_1^* y_1 + x_2^* y_2 + \cdots + x_n^* y_n \tag{8}$$

When $\mathbf{x}$ and $\mathbf{y}$ are real vectors, the inner product becomes

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \tag{9}$$

Bilinear Form. A scalar homogeneous expression containing the product of the elements of vectors $\mathbf{x}$ and $\mathbf{y}$ is called a *bilinear form* in the variables x_{ij} and y_{ij}. When they are both of order n, the most general bilinear form in $\mathbf{x}$ and $\mathbf{y}$ is

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) = {} & a_{11}x_1y_1 + a_{12}x_1y_2 + \cdots + a_{1n}x_1y_n \\ & + a_{21}x_2y_1 + a_{22}x_2y_2 + \cdots + a_{2n}x_2y_n \\ & + \cdots \\ & + a_{n1}x_ny_1 + a_{n2}x_ny_2 + \cdots + a_{nn}x_ny_n \end{aligned} \tag{10}$$

This can be written more compactly as

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j \\ &= [x_1x_2\cdots x_n]\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} \\ &= \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{y} = \langle \mathbf{x}, \mathbf{A}\mathbf{y} \rangle \end{aligned} \tag{11}$$

The matrix $\mathbf{A}$ is called the coefficient matrix of the bilinear form, and the rank of $\mathbf{A}$ is called the rank of the bilinear form. A bilinear form is called *symmetric* if the matrix $\mathbf{A}$ is symmetric.

Quadratic Form. A quadratic form V is a real homogeneous polynomial in the real variables $x_1, x_2, \ldots, x_n$ of the form

$$V = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \tag{12}$$

where all a_{ij} are real. This is the special case of Eq. (11) where $\mathbf{x} = \mathbf{y}$. The quadratic form V can therefore be expressed as the inner product

$$V(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle \tag{13}$$

A homogeneous polynomial can always be expressed in terms of a symmetric matrix $\mathbf{A}$. In the expansion of Eq. (12) the cross-product terms ($i \neq j$) all have the form $(a_{ij} + a_{ji})x_ix_j$. Choosing $a_{ij} = a_{ji}$ makes the matrix $\mathbf{A}$ for the quadratic form symmetric. This is illustrated in Eq. (14).

$$\begin{aligned} V(x) &= x_1^2 - 4x_2^2 + 5x_3^2 + 6x_1x_2 - 20x_2x_3 \\ &= x^{\mathrm{T}}\begin{bmatrix} 1 & 3 & 0 \\ 3 & -4 & -10 \\ 0 & -10 & 5 \end{bmatrix}x = x^{\mathrm{T}}Ax \end{aligned} \tag{14}$$

The rank of the matrix $\mathbf{A}$ is called the rank of the quadratic form. If the rank of $\mathbf{A}$ is $r < n$, the quadratic form is singular. If the rank of $\mathbf{A}$ is n, the quadratic form is nonsingular. A nonsymmetric matrix can be converted into an equivlent symmetric matrix by replacing all sets of elements a_{ij} and a_{ji} by the average value $(a_{ij} + a_{ji})/2$. The principal diagonal will be preserved.

Length of a Vector. The length of a vector $\mathbf{x}$ is called the Euclidean norm and is denoted by $\|\mathbf{x}\|$. It is defined as the square root of the inner product $\langle \mathbf{x}, \mathbf{x} \rangle$. For real vectors it is given by

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \tag{15}$$

A vector can be normalized so that its length is unity. In that case it is called a *unit vector.* The unit vector may be denoted by $\hat{\mathbf{x}}$ and is obtained by dividing each element of $\mathbf{x}$ by $\|\mathbf{x}\|$:

$$\hat{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|} \tag{16}$$

Principal Minor. A principal minor of a matrix $\mathbf{A}$ is obtained by deleting any row(s) and the same numbered col-

umn(s). The diagonal elements of a principal minor of **A** are therefore also diagonal elements of **A**. The determinant |A| is classified as a principal minor, with no rows and columns deleted. The number of principal minors can be determined from a Pascal triangle. The number of principal minors for a matrix **A** of order n is: 1 for $n = 1$, 3 for $n = 2$, 7 for $n = 3$, and 15 for $n = 4$, ...

Leading Principal Minor. There are n leading principal minors, formed by all the square arrays within **A** that contain a_{11}. Starting with $[a_{11}]$, the next square array is formed by including the next row and column. This process is continued until the matrix **A** is obtained. The leading principal minors are given by

$$\Delta_1 = |a_{11}| \qquad \Delta_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$
$$\Delta_{123} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \cdots \quad \Delta_{12\cdots n} = |\mathbf{A}| \tag{17}$$

The subscripts of Δ are the rows and columns of **A** used to form the minor. The minor contains the common elements of these rows and columns.

Definiteness and Semidefiniteness. The sign definiteness for a scalar function of a vector, such as $V(\mathbf{x})$, is defined for a spherical region S about the origin described by $\|\mathbf{x}\| \leq K$ (a constant equal to the radius of S). The function $V(\mathbf{x})$ and all $\partial V(\mathbf{x})/\partial x_i$ for $i = 1, 2, \ldots, n$ must be continuous within S. The definiteness of a quadratic form is determined by analyzing only the symmetric **A** matrix. If **A** is given as a nonsymmetric matrix, it must first be converted to a symmetric matrix.

Positive Definite (PD). A scalar function, such as the quadratic form $V(\mathbf{x}) = \langle \mathbf{x}, \mathbf{Ax} \rangle$, is called positive definite when $V(\mathbf{x}) = 0$ for $\mathbf{x} = \mathbf{0}$ and $V(\mathbf{x}) > 0$ for all other $\|\mathbf{x}\| \leq K$. The positive definite condition requires that $|\mathbf{A}| \neq \mathbf{0}$; thus, the rank of **A** is equal to n. The rank of **A** may be determined by putting it in hermite normal form. When a real quadratic form $V(\mathbf{x}) = \mathbf{x}^T\mathbf{Ax}$ is positive definite, then the matrix **A** is also called positive definite. The matrix **A** has the property that it is positive definite *iff* there exists a nonsingular matrix **H** such that $\mathbf{A} = \mathbf{H}^T\mathbf{H}$.

The definiteness of the matrix **A** can be determined by reducing this matrix to diagonal form by means of a congruent transformation such that the diagonal matrix $\mathbf{B} = \mathbf{P}^T\mathbf{AP}$. The matrix **P** may be formed by using the following elementary transformations:

1. Interchanging the ith and jth rows and interchanging the ith and jth columns.
2. Multiplying the ith row and column by a nonzero scalar k.
3. Addition to the elements of the ith row of the corresponding elements of the jth row multiplied by a scalar k and addition to the ith column of the jth column multiplied by k.

If the principal diagonal of the matrix **B** contains only positive, nonzero elements, then the matrix **A** is positive definite.

Another method of diagonalizing the **A** matrix is to determine a modal matrix **T** (8). In that case the diagonalized matrix contains the eigenvalues obtained from $|\lambda\mathbf{I} - \mathbf{A}| = 0$. In order for the matrix **A** to be positive definite, all the eigenvalues must be positive.

An alternate method of determining positive definiteness is to calculate *all* the leading principal minors of the matrix **A**. If all the leading principal minors of **A** are positive, the real quadratic form is positive definite. (This is sometimes called the Sylvester theorem.) Note that when the leading principal minors are all positive, then all the other principal minors are also positive.

Positive Semidefinite (PSD). The quadratic form $V(\mathbf{x})$ is called positive semidefinite when $V(\mathbf{x}) = 0$ for $\mathbf{x} = \mathbf{0}$ and $V(\mathbf{x}) \geq 0$ for all $\|\mathbf{x}\| \leq K$. The function $V(\mathbf{x})$ is permitted to equal zero at points in S other than the origin (i.e., for some $\mathbf{x} \neq \mathbf{0}$), but it may not be negative. In this case $\|\mathbf{A}\| = 0$, the rank of **A** is less than n, and all of the eigenvalues of **A** are positive. The positive semidefinite quadratic form can therefore be reduced to the form $y_1^2 + y_2^2 + \cdots + y_r^2$, where $r < n$. When $V(\mathbf{x})$ is positive semidefinite, then the matrix **A** is also called positive semidefinite. The matrix **A** is positive semidefinite *iff* all its characteristic values are ≥ 0. In order to be positive semidefinite, *all* the principal minors (9) must be nonnegative.

Negative Definite (ND) and Negative Semidefinite (NSD). The definitions of negative definite and negative semidefinite follow directly from the definitions above when the inequalities are applied to $-\mathbf{A}$. If a matrix **A** does not satisfy the conditions for positive definiteness or positive semidefiniteness, then $-\mathbf{A}$ is checked for these conditions. If $-\mathbf{A}$ satisfies the positive definite or positive semidefinite conditions, then **A** is said to be negative definite or negative semidefinite, respectively.

Indefinite. A scalar function $V(\mathbf{x})$ is *indefinite* if it assumes both positive and negative values within the region S described by $\|\mathbf{x}\| < K$. If **A** is not positive definite, positive semidefinite, negative definite, or negative semidefinite, then it is said to be indefinite. Note that when **A**, in general form, has both positive and negative elements along the principal diagonal, it is indefinite. When **A** is transformed to a diagonal form, it is indefinite if some of the diagonal elements are positive and some are negative. This means that some of the eigenvalues are positive and some are negative.

Example 1. Use the principal minors to determine the definiteness of

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \tag{18}$$

The leading principal minors are evaluated to check for positive definiteness:

$$\Delta_1 = 1, \qquad \Delta_{12} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0, \qquad \Delta_{123} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \Delta = 0$$

The conditions for positive definiteness are not satisfied. Next the remaining principal minors are evaluated to check for positive semidefiniteness.

$$\Delta_{13} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1, \qquad \Delta_{23} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1$$
$$\Delta_2 = 1, \qquad \Delta_3 = 0$$

Since the principal minors are not all nonnegative, the matrix **A** is not positive semidefinite. Similarly, **A** is not negative definite or negative semidefinite; therefore **A** is indefinite.

Example 2. Transform **A** in Eq. (18) to the diagonal Jordan form Λ and check its definiteness. The characteristic equation is

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 1 & -1 & -1 \\ -1 & \lambda - 1 & -1 \\ -1 & -1 & \lambda \end{vmatrix} = \lambda^3 - 2\lambda^2 - 2\lambda = 0 \qquad (19)$$

The eigenvalues are $\lambda_1 = 2.732$, $\lambda_2 = -0.732$, and $\lambda_3 = 0$. Thus the matrix **A** is indefinite since

$$\Lambda = \begin{bmatrix} 2.732 & 0 & 0 \\ 0 & -0.732 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

LINEARIZATION (JACOBIAN MATRIX) (3–5,7,10)

A linear system with no forcing function (an autonomous system) and with $|\mathbf{A}| \neq 0$ has only one equilibrium point $\mathbf{x}_0$. A nonlinear system, on the other hand, may have more than one equilibrium point. This is easily illustrated by considering the unity-feedback angular position-control system shown in Fig. 2. The feedback action is provided by synchros which generate the actuating signal $e = \sin(\theta_i - \theta_o)$. With no input, $\theta_i = 0$, the differential equation of the system is

$$\ddot{\theta}_o + a\dot{\theta}_o + K \sin \theta_o = 0 \qquad (20)$$

This is obviously a nonlinear differential equation because of the term $\sin \theta_o$. With the phase variables $x_1 = \theta_o$ and $x_2 = \dot{x}_1 = \dot{\theta}_o$, the corresponding state equations are

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = -K \sin x_1 - a x_2 \qquad (21)$$

The slope of the trajectories in the phase plane (25) is obtained from

$$N = \frac{\dot{x}_2}{\dot{x}_1} = \frac{-K \sin x_1 - a x_2}{x_2} \qquad (22)$$

In the general case the unforced state equation is $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. Since equilibrium points $\mathbf{x}_0$ exist at $\dot{\boldsymbol{x}} = f(x_0) = 0$, the singularities are $x_2 = 0$ and $x_1 = k\pi$, where k is an integer. The system therefore has multiple equilibrium points.

In a small neighborhood about each of the equilibrium points, a nonlinear system behaves like a linear system. The states can therefore be written as $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}^*$, where $\mathbf{x}^*$ represents the perturbation or *state deviation* from the equilibrium point $\mathbf{x}_0$. Each of the elements of $\mathbf{f}(\mathbf{x})$ can be expanded

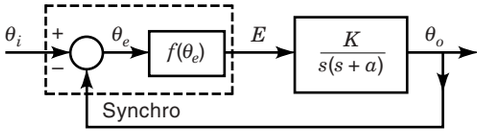

Figure 2. A nonlinear feedback control system.

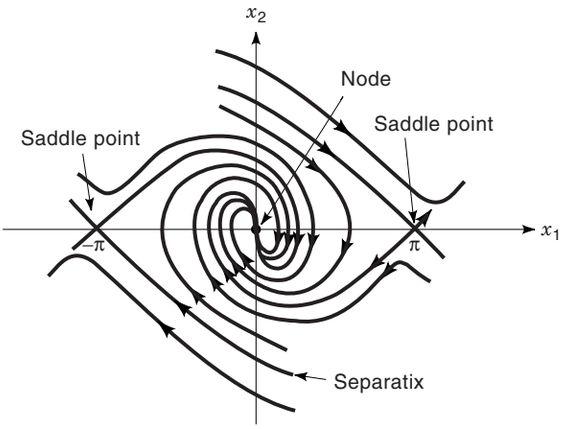

Figure 3. Phase portrait for Eq. (20).

in a Taylor series about one of the equilibrium points $\mathbf{x}_0$. Assuming that $\mathbf{x}^*$ is restricted to a small neighborhood of the equilibrium point, the higher-order terms in the Taylor series may be neglected. Thus, the resulting linear variational state equation is

$$\dot{\mathbf{x}}^* = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ & \cdots & & \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}_{\mathbf{x}=\mathbf{x}_0} \mathbf{x}^* = \mathbf{J}_x \mathbf{x}^* \qquad (23)$$

where f_i is the ith row of $\mathbf{f}(\mathbf{x})$ and $\mathbf{J}_x = \partial \mathbf{f} / \partial \mathbf{x}^{\mathrm{T}}$ is called the *Jacobian* matrix and is evaluated at $\mathbf{x}_0$.

For the system of Eq. (21) the motion about the equilibrium point $x_1 = x_2 = 0$ is represented by

$$\dot{\mathbf{x}}^* = \begin{bmatrix} \dot{x}_1^* \\ \dot{x}_2^* \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -K & -a \end{bmatrix} \begin{bmatrix} x_1^* \\ x_1^* \end{bmatrix} = \mathbf{J}_x \mathbf{x}^* \qquad (24)$$

For these linearized equations the eigenvalues are $\lambda_{1,2} = -a/2 \pm \sqrt{(a/2)^2 - K}$. This equilibrium point is stable and is either a node or a focus, depending upon the magnitudes of a and K. A node denotes an overdamped response and a focus denotes an underdamped response about the equilibrium point. For motion about the equilibrium point $x_1 = \pi$, $x_2 = 0$, the state equation is

$$\dot{\mathbf{x}}^* = \begin{bmatrix} 0 & 1 \\ K & -a \end{bmatrix} \mathbf{x}^* \qquad (25)$$

The eigenvalues of $\mathbf{J}_x$ are $\lambda_{1,2} = -a/2 \pm \sqrt{(a/2)^2 + K}$. Thus, one eigenvalue is positive and the other is negative, and the equilibrium point represents an unstable saddle point. The motion around the saddle point is considered unstable since every point on all trajectories, except on the two separatrices, moves away from this equilibrium point. A phase-plane portrait for this system can be obtained by the method of isoclines (25). From Eq. (22) the isocline equation is $x_2 = -K \sin x_1/(N + a)$. The phase portrait is shown in Fig. 3. Note that the linearized

equations are applicable only in the neighborhood of the singular points. Thus, they describe stability in the small.

In analyzing the system performance in the vicinity of equilibrium, it is usually convenient to translate the origin of the state space to that point. This is done by inserting $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}^*$ into the original equations. With the origin at $\mathbf{x}_0$, the variations from this point are described by $\mathbf{x}^*$.

SECOND METHOD OF LYAPUNOV

The second, or direct, method of Lyapunov provides a means for determining the stability of a system without explicitly solving for the trajectories in the state space. This is in contrast to the first method of Lyapunov, which requires the determination of the eigenvalues from the linearized equations about an equilibrium point. The second method is applicble for determining the behavior of higher-order systems which may be forced or unforced, linear or nonlinear, time-invariant or time-varying, and deterministic or stochastic. Solution of the differential equation is not required. The procedure requires the selection of a scalar "energy" function $V(\mathbf{x})$, which is tested for the conditions that indicate stability. When $V(\mathbf{x})$ successfully meets these conditions, it is called a *Lyapunov function.* The principal difficulty in applying the method is in formulating a correct Lyapunov function because, for asymptotically stable systems, the failure of one function to meet the stability conditions does not mean that a true Lyapunov function does not exist. This difficulty is compounded by the fact that the Lyapunov function is not unique. Nevertheless, there is much interest in the second method (10,11).

In order to show a simple example of a Lyapunov function, consider the system of Fig. 2 represented by Eq. (20) which has multiple equilibrium points at $\dot{\theta}_o = 0$, $\theta_o = n\pi$, where n is an integer. The proposed function is the sum of the kinetic and potential stored energies, given by

$$V(\theta_o, \dot{\theta}_o) = \tfrac{1}{2}\dot{\theta}_o^2 + K(1 - \cos\theta_o) \tag{26}$$

This function is positive for all values of θ_o and $\dot{\theta}_o$, except at the equilibrium points, where it is equal to zero. The rate of change of this energy function along any phase-plane trajectory is obtained by differentiating Eq. (26) and using Eq. (20):

$$\dot{V}(\theta_o, \dot{\theta}_o) = \dot{\theta}_o\ddot{\theta}_o + K(\sin\theta_o)\dot{\theta}_o = -a\dot{\theta}_o^2 \tag{27}$$

The value of $\dot{V}$ is negative along any trajectory for all values of $\dot{\theta}_o$, *except* $\dot{\theta}_o = 0$. Note that the slope $d\dot{\theta}_o/d\theta_o$ of the phase-plane trajectories [see Eq. (22)] is infinite along the line represented by $\dot{\theta}_o = 0$, except at the equilibrium points $\theta_o = n\pi$. Thus, the line $\dot{\theta}_o = 0$ does not represent an equilibrium, except at $\theta_o = n\pi$, and it is not a trajectory in the state plane. Since the energy stored in the system is continuously decreasing at all points except at the equilibrium points, the equilibrium at the origin and at even multiples of $\theta_o = n\pi$ is asymptotically stable. This $\dot{V}$ is NSD.

This example demonstrates that the total system energy may be used as the Lyapunov function. When the equations of a large system are given in mathematical form, it is usually difficult to define the energy of the system. Thus, alternate Lyapunov functions must be obtained. For any system of order n, a positive, constant value of the proper positive definite Lyapunov function $V(\mathbf{x})$ represents a closed surface, a hyperellipsoid, in the state space with its center at the origin. The entire state space is filled with such nonintersecting closed surfaces, each representing a different positive value of $V(\mathbf{x})$. When $\dot{V}(\mathbf{x})$ is negative for all points in the state space, it means that all trajectories cross the closed surfaces from the outside to the inside and eventually converge at the equilibrium point at the origin.

A qualitatively correct Lyapunov function is shown in Fig. 4 for a second-order system. The function $V(x_1, x_2) = x_1^2 + x_2^2$ is positive definite and is represented by the paraboloid surface shown. The value $V(x_1, x_2) = k_i$ (a constant) is represented by the intersection of the surface $V(x_1, x_2)$ and the plane $z = k_i$. The projection of this intersection on the x_1, x_2 plane is a closed curve, an oval, around the origin. There is a family of such closed curves in the x_1, x_2 plane for different values of k_i. The value $V(x_1, x_2) = 0$ is the point at the origin; it is the innermost curve of the family of curves representing different levels on the paraboloid.

The gradient vector of $V(\mathbf{x})$ is defined by

$$\text{gradient } V(\mathbf{x}) = \nabla V(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial V(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial V(\mathbf{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial V(\mathbf{x})}{\partial x_n} \end{bmatrix} \tag{28}$$

The time derivative $\dot{V}(x)$ along any trajectory is

$$\begin{aligned} \dot{V}(x) &= \frac{\partial V(\mathbf{x})}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial V(\mathbf{x})}{\partial x_2}\frac{dx_2}{dt} + \cdots + \frac{\partial V(\mathbf{x})}{\partial x_n}\frac{dx_n}{dt} \\ &= [\nabla V(\mathbf{x})]^T\dot{\mathbf{x}} = \langle \nabla V, \dot{\mathbf{x}} \rangle \end{aligned} \tag{29}$$

It is important to note that $\dot{V}(\mathbf{x})$ can be evaluated without knowing the solution of the system state equation $\dot{\mathbf{x}} = f(\mathbf{x})$. The gradient $\nabla V(x_1, x_2)$ describes the steepness between adjacent levels of $V(x_1, x_2)$. The function $V(x_1, x_2)$ is negative definite in Fig. 4—except at the origin, where it is equal to zero. The state-plane trajectory crosses the ovals for successively smaller values of $V(\mathbf{x})$. Therefore, the system is asymptoti-

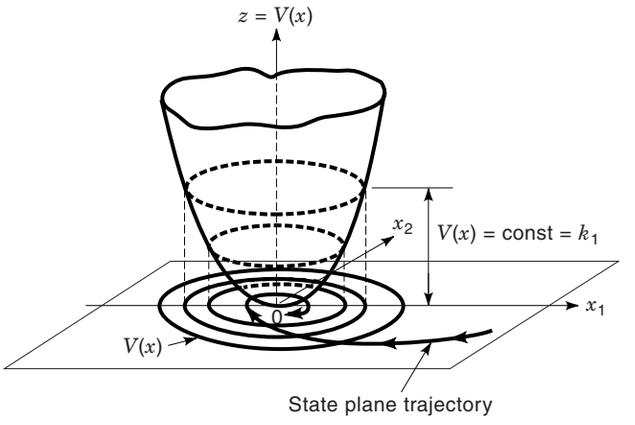

Figure 4. A positive definite function $V(x_1, x_2)$ and projections on the x_1, x_2 plane.

cally stable, and $V(\mathbf{x})$ is a proper Lyapunov function. The concepts described above may be summarized in the following theorem, which provides sufficient, but not necessary, conditions for stability.

Theorem 1. Lyapunov asymptotic stability. A system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, where $\mathbf{f}(\mathbf{0}, t) = \mathbf{0}$ for all t, is asymptotically stable in the vicinity of the equilibrium point at the origin if there exists a scalar function $V(\mathbf{x})$ such that:

1. $V(\mathbf{x})$ is continuous and has continuous first partial derivatives in a region S around the origin.
2. $V(\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0}$. [$V(\mathbf{x})$ is PD.]
3. $V(\mathbf{0}) = 0$.
4. $\dot{V}(\mathbf{x}) < 0$ for $\mathbf{x} \neq \mathbf{0}$. [$\dot{V}(x)$ is ND.]

Conditions 1–3 ensure that $V(\mathbf{x})$ is positive definite. Therefore, $V(\mathbf{x}) = k$ is a closed surface within the region S. Condition 4 means that $\dot{V}(x)$ is negative definite, and thus any trajectory in S crosses through the surface $V(\mathbf{x}) = k$ from the outside to the inside for all values of k. Therefore, the trajectory converges on the origin where $V(\mathbf{0}) = 0$.

The condition that $\dot{V}(\mathbf{x}) < 0$ for $x \neq 0$ can be relaxed in Theorem 1 under the proper conditions. Condition 4 can be changed to $\dot{V}(\mathbf{x}) \leq 0$; that is, $\dot{V}(\mathbf{x})$ is negative semidefinite. This relaxed condition is sufficient, provided that $\dot{V}(\mathbf{x})$ is not equal to zero at any solution of the original differential equation except at the equilibrium point at the origin. A test for this condition is to insert the solution of $\dot{V}(\mathbf{x}) = 0$ into the state equation $\dot{\mathbf{x}} = f(\mathbf{x})$ to verify that it is satisfied only at the equilibrium point. Also, if it can be shown that no trajectory can stay forever at the points or on the line, other than the origin, at which $\dot{V} = 0$, then the origin is asymptotically stable. This is the case for the system of Fig. 2 as described at the beginning of this section. For linear systems there is only one equilibrium point which is at the origin; therefore it is sufficient for $\dot{V}(\mathbf{x})$ to be negative semidefinite. Theorem 1 may also be extended so that it is applicable to the entire state space. In that case the system is said to have global stability, or stability in the large. Including these conditions results in the following theorem:

Theorem 2. Lyapunov global asymptotic stability. A system is globally asymptotically stable if there is only one stable equilibrium point and there exists a scalar function $V(\mathbf{x})$ such that:

1. $V(\mathbf{x})$ is continuous and has continuous first partial derivatives in the entire state space.
2. $V(\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0}$.
3. $V(\mathbf{0}) = 0$.
4. $V(\mathbf{x}) \to \infty$ at $\|\mathbf{x}\| \to \infty$.
5. $\dot{V}(\mathbf{x}) \leq 0$.
6. Either $\dot{V}(\mathbf{x}) \neq \mathbf{0}$, except at $\mathbf{x} = 0$, or any locus in the state space where $\dot{V}(\mathbf{x}) = 0$ is not a trajectory of the system.

Conditions 1–3 ensure that $V(\mathbf{x})$ is positive definite. Condition 4 is satisfied when $V(\mathbf{x})$ is positive definite (i.e., it is closed) in the entire state space. When $V(\mathbf{x})$ goes to infinity as any $x_i \to \infty$, then $V(\mathbf{x}) = k_i$ is a closed curve for any k_i. Conditions 5 and 6 mean that $V(\mathbf{x})$ is continuously decreasing along any trajectory in the entire plane and ensures that the system is asymptotically stable. In order to check for global stability, it is necessary to select $V(\mathbf{x})$ so that conditions 1–4 are all satisfied.

Finding a proper Lyapunov function $V(\mathbf{x})$ means that the system is stable, but this is just a sufficient and not a necessary condition for stability. The fact that a function $V(\mathbf{x})$ has not been found does not mean that it does not exist and that the system is not stable.

Application of the Lyapunov Method to Linear Systems

The second method of Lyapunov is applicable to time-varying and nonlinear systems. However, there is no simple general method of developing the Lyapunov function. Methods have been developed for many such systems, and the literature in this area (4) is extensive. The remaining applications in this chapter are restricted to linear systems. Since linear systems may have only one equilibrium point, the stability or instability is necessarily global in nature. The Routh–Hurwitz stability criterion is available for determining the stability of linear systems. However, the necessity of obtaining the characteristic polynomial can be a disadvantage for higher-order systems. Thus, the following material is presented to develop familiarity with the more general second method of viewing stability. Confidence in the second method of Lyapunov can be developed by showing that the results are identical to those obtained with the Routh–Hurwitz method.

The approach presented is to select a $V(\mathbf{x})$ which is positive definite and to evaluate $\dot{V}(\mathbf{x})$. The coefficients of $V(\mathbf{x})$ and those restraints on the system parameters are then determined which make $\dot{V}(\mathbf{x})$ negative definite or negative semidefinite. Consider the linear, unity feedback system presented in Fig. 5 with $r(t) = 0$.

Example

$$G(s) = \frac{K}{s(s+a)}, \qquad a > 0 \tag{30}$$

The differential equation for the actuating signal, when $r = 0$, is

$$\ddot{e} + a\dot{e} + Ke = 0 \tag{31}$$

When phase variables with $x_1 = e$ are used, the state equations are

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = -Kx_1 - ax_2 \tag{32}$$

A simple Lyapunov function which is positive definite is

$$V(\mathbf{x}) = \frac{1}{2}p_1x_1^2 + \frac{1}{2}p_2x_2^2 = \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x} \tag{33}$$

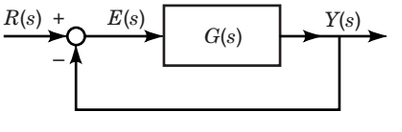

Figure 5. Unity-feedback linear system.

where $p_1 > 0$ and $p_2 > 0$. Its derivative is

$$\dot{V}(\mathbf{x}) = p_1 x_1 \dot{x}_1 + p_2 x_2 \dot{x}_2 = p_1 x_1 x_2 - p_2 K x_1 x_2 - a p_2 x_2^2 = \mathbf{x}^{\mathrm{T}} \begin{bmatrix} 0 & \dfrac{p_1 - p_2 K}{2} \\ \dfrac{p_1 - p_2 K}{2} & -a p_2 \end{bmatrix} \mathbf{x} = -\mathbf{x}^{\mathrm{T}} \mathbf{N} \mathbf{x} \tag{34}$$

The function $\dot{V}(\mathbf{x})$ is always negative semidefinite if $-\mathbf{N}$ is negative semidefinite. The matrix $-\mathbf{N}$ is negative semidefinite if *all* the principal minors of $\mathbf{N}$ are positive or zero.

$$\Delta_2 = a p_2 \geq 0 \tag{35}$$

$$\Delta_{12} = -\frac{(p_1 - p_2 K)^2}{4} \geq 0 \tag{36}$$

Equation (35) is satisfied since it is required that both a and p_2 be greater than zero. Since the left side of Eq. (36) cannot be positive, only the equality condition $p_1 - p_2 K = 0$ can be satisfied, which yields $p_1 = p_2 K$. Since $p_1 > 0$ and $p_2 > 0$, this requires that $K > 0$. Then Eq. (34) yields $\dot{V}(\mathbf{x}) = -a p_2 x_2^2$, which is negative semidefinite. The Lyapunov global asymptotic stability theorem (Theorem 2) is satisfied since, by condition 4, $V(\mathbf{x}) \to \infty$ as $\|\mathbf{x}\| \to \infty$. The condition $\dot{V}(\mathbf{x}) = 0$ holds along the x_1 axis, where $x_2 = 0$ and x_1 has any value. A way of showing that $\dot{V}(\mathbf{x})$ being negative semidefinite is sufficient for global asymptotic stability is to show that the x_1 axis is not a trajectory of the system state equations [see Eq. (32)]. The first equation yields $\dot{x}_1 = 0$ or $x_1 = c$. The x_1 axis can be a trajectory only if $x_2 = 0$ and $\dot{x}_2 = 0$. But the second state equation yields $\dot{x}_2 = -K x_1 = -K c = 0$. This is a contradiction since it requires $x_1 = c = 0$. Therefore, the x_1 axis is not a trajectory, and the system is asymptotically stable. Using Lyapunov's second method, the necessary condition for stability is shown above to be $K > 0$. This result is readily recognized as being correct, either from the Routh stability criterion or from the root locus.

Lur'e Method

Another design approach is the use of a procedure developed by Lur'e (cited in Ref. 12). Consider the unexcited system represented by the state equation in which $\mathbf{A}$ is of order n:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{37}$$

The quadratic Lyapunov function $V(\mathbf{x})$ is expressed in terms of the symmetric matrix $\mathbf{P}$ by

$$V(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} \mathbf{P} \mathbf{x} \tag{38}$$

The time derivative of $V(\mathbf{x})$ is

$$\dot{V}(\mathbf{x}) = \dot{\mathbf{x}}^{\mathrm{T}} \mathbf{P} \mathbf{x} + \mathbf{x}^{\mathrm{T}} \mathbf{P} \dot{\mathbf{x}} = \mathbf{x}^{\mathrm{T}} (\mathbf{A}^{\mathrm{T}} \mathbf{P} + \mathbf{P}\mathbf{A}) \mathbf{x} = -\mathbf{x}^{\mathrm{T}} \mathbf{N} \mathbf{x} \tag{39}$$

where the symmetric matrix $\mathbf{N}$ is given by

$$\mathbf{N} = -(\mathbf{A}^{\mathrm{T}} \mathbf{P} + \mathbf{P}\mathbf{A}) = -2(\mathbf{P}\mathbf{A})_{\mathrm{sym}} \tag{40}$$

The following procedure is used:

Step 1. Select an arbitrary symmetric matrix $\mathbf{N}$ of order n which is either positive definite or positive semidefinite. For example, a simple diagonal matrix such as $\mathbf{N} = \mathbf{I}$ or $\mathbf{N} = 2\mathbf{I}$ is positive definite. A positive semidefinite matrix $\mathbf{N}$ can be chosen which contains all zero elements except one positive element in any position along the principal diagonal.

Step 2. Determine the elements of $\mathbf{P}$ by equating terms in Eq. (40). Since $\mathbf{P}$ is symmetric, this requires the solution of $n(n + 1)/2$ equations.

Step 3. Use the Sylvester theorem to determine the definiteness of $\mathbf{P}$.

Step 4. Since $\mathbf{N}$ is selected as positive definite (or positive semidefinite), the necessary and sufficient condition for asymptotic stability (instability) is that $\mathbf{P}$ be positive definite (negative definite).

The Sylvester conditions for the positive definiteness of $\mathbf{P}$ are the same as the Routh–Hurwitz stability conditions. The equivalence is derived in Ref. 12. Linear systems which are asymptotically stable are globally asymptotically stable.

Example. Examine the conditions of asymptotic stability for the second-order dynamic system

$$\ddot{x} + a_1 \dot{x} + a_0 x = 0 \tag{41}$$

With phase variables the coefficient matrix is

$$A = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} \tag{42}$$

The symmetric matrix $\mathbf{P}$ and a chosen positive semidefinite matrix $\mathbf{N}$ are

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}, \qquad \mathbf{N} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \tag{43}$$

The identity in Eq. (40) yields three simultaneous equations: The solution of these equations yields

$$\begin{aligned} n_{11} &= 2 = 2a_0 p_{12} \\ n_{12} &= 0 = -p_{11} + a_1 p_{12} + a_0 p_{22} \\ n_{22} &= 0 = -2p_{12} + 2a_1 p_{22} \\ p_{12} &= \frac{1}{a_0}, \qquad p_{22} = \frac{1}{a_0 a_1}, \qquad p_{11} = \frac{a_0 + a_1^2}{a_0 a_1} \end{aligned} \tag{44}$$

The necessary conditions for $\mathbf{P}$ to be positive definite are obtained by applying the Sylvester theorem:

$$p_{11} = \frac{a_0 + a_1^2}{a_0 a_1} > 0, \qquad p_{11} p_{22} - p_{12}^2 = \frac{a_0}{a_0^2 a_1^2} > 0 \tag{45}$$

The second equation requires that $a_0 > 0$. Using this condition in the first equation produces the necessary condition $a_1 > 0$. These are obviously the same conditions that are obtained from the Routh–Hurwitz conditions for stability, as shown in a later section.

Krasovskii (cited in Ref. 4) has shown that a similar approach may be used with a nonlinear system. However, only sufficient conditions for local asymptotic stability in the vicinity of an equilibrium point may be obtained. For the Lyapu-

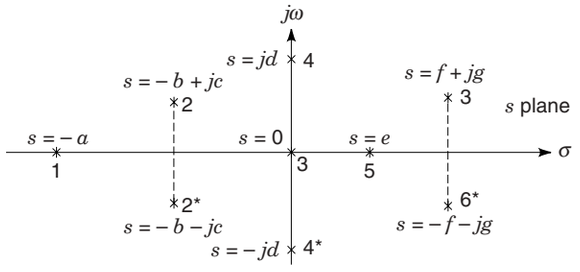

Figure 6. Location of poles in the s plane. (Numbers are used to identify the poles.)

nov function $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$, the time derivative is $\dot{V}(\mathbf{x}) = \mathbf{x}^T\mathbf{N}\mathbf{x}$, where

$$-\mathbf{N} = \mathbf{J}_\mathbf{x}^T\mathbf{P} + \mathbf{P}\mathbf{J}_\mathbf{x} \tag{46}$$

The matrix $\mathbf{J}_\mathbf{x}$ is the Jacobian evaluated at the equilibrium point. Selecting $\mathbf{P} = \mathbf{I}$ may often lead to a successful determination of the conditions for asymptotic stability in the vicinity of the equilibrium point. This extension to nonlinear systems shows the generality of the method.

LOCATION OF POLES AND STABILITY

The dynamic equations of a system are often developed in linearized form. However, if they are nonlinear, it is possible to obtain a linearized approximation as described in the section entitled "Linearization (Jacobian Matrix)." In terms of the Laplace transform variable s, the response transform $Y(s)$ for the linear representation can be expressed, in general, as the ratio of two polynomials. The response transform $Y(s)$ has the form

$$\begin{aligned} Y(s) &= \frac{P(s)}{Q(s)}R(s) = \frac{a_w s^w + a_{w-1}s^{w-1} + \cdots + a_1 s + a_0}{s^n + b_{n-1}s^{n-1} + \cdots + b_1 s + b_0}R(s) \\ &= \frac{P(s)}{(s-s_1)(s-s_2)\cdots(s-s_k)\cdots(s-s_n)}R(s) \end{aligned} \tag{47}$$

where $R(s)$ is the system input and the values of the poles s_1, s_2, . . ., s_n may be real or occur in conjugate complex pairs.

The stability and the corresponding response of a system can be determined from the locations of the poles of the response transform $Y(s)$ in the s plane. The possible positions of the poles are shown in Fig. 6, and the responses associated with the poles are given in Table 1. These poles are the roots of the characteristic equation $Q(s) = 0$.

Table 1. Relation of Response to Location of Poles

Position of Pole	Form of Response	Characteristics
1	Ae^{-at}	Damped exponential
2–2*	$Ae^{-bt}\sin(ct + \phi)$	Exponentially damped sinusoid
3	A	Constant
4–4*	$A\sin(dt + \phi)$	Constant-sinusoid
5	Ae^{et}	Increasing exponential (unstable)
6–6*	$Ae^{ft}\sin(gt + \phi)$	Exponentially increasing sinusoid (unstable)

Poles of the response transform located at the origin or on the imaginary axis that are not contributed by the forcing function $R(s)$ are undesirable in a control system. Poles in the right half-plane result in transient terms that increase with time. Such performance characterizes an *unstable system*. Therefore, poles in the right-half s plane are not permissible.

Routh Stability Criterion

The stability characteristic of a linear time-invariant system is determined from the system's characteristic equation (13). A system is asymptotically stable if and only if all roots of the characteristic equation lie to the left of the imaginary axis. Routh's stability criterion provides a means for determining stability without evaluating the roots of this equation. The response transform $Y(s)$ has the general form given by Eq. (47), where $R(s)$ is the driving transform. Although the inverse Laplace transform of $Y(s)$ can be performed, the polynomial $Q(s)$ must first be factored. Computer and calculator programs are readily available for obtaining the roots of a polynomial (18). Figure 6 and Table 1 show that stability of the response $y(t)$ requires that all zeros of $Q(s)$ have negative real parts. Since it is usually not necessary to find the exact solution when the response is unstable, a simple procedure to determine the existence of zeros with positive real parts is needed. If such zeros of $Q(s)$ with positive real parts are found, the system is unstable and must be modified. Routh's criterion is a simple method of determining the number of zeros with positive real parts without actually solving for the zeros of $Q(s)$. Note that zeros of $Q(s)$ are poles of $Y(s)$ (15,16).

The characteristic equation is

$$Q(s) = b_n s^n + b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \cdots + b_1 s + b_0 = 0 \tag{48}$$

If the b_0 term is zero, divide by s to obtain the equation in the form of Eq. (48). The b's are real coefficients, and all powers of s from s^n to s^0 must be present in the characteristic equation. A necessary but not sufficient condition for stable roots is that all the coefficients in Eq. (48) be positive. If any coefficients other than b_0 are zero or if the coefficients do not all have the same sign, then there are pure imaginary roots or roots with positive real parts and the system is unstable. It is therefore unnecessary to continue if only stability or instability is to be determined. When all the coefficients are present and positive, the system may or may not be stable because there still may be roots on the imaginary axis or in the right-half s plane. Routh's criterion is mainly used to determine stability. In special situations it may be necessary to determine the actual number of roots in the right-half s plane. For these situations the procedure described in this section can be used.

The coefficients of the characteristic equation are arranged in the pattern shown in the first two rows of the following Routhian array. These coefficients are then used to evaluate the rest of the constants to complete the array.

$$\begin{array}{lllll} s^n \mid b_n & b_{n-2} & b_{n-4} & b_{n-6} & \cdots \\ s^{n-1} \mid b_{n-1} & b_{n-3} & b_{n-5} & b_{n-7} & \cdots \\ s^{n-2} \mid c_1 & c_2 & c_3 & \cdots & \\ s^{n-3} \mid d_1 & d_2 & \cdots & & \\ \cdots \mid \cdots & \cdots & & & \\ s^1 \mid j_1 & & & & \\ s^0 \mid k_1 & & & & \end{array} \tag{49}$$

The constants c_1, c_2, c_3, and so on, in the third row are evaluated as follows:

$$c_1 = \frac{b_{n-1}b_{n-2} - b_n b_{n-3}}{b_{n-1}} \tag{50}$$

$$c_2 = \frac{b_{n-1}b_{n-4} - b_n b_{n-5}}{b_{n-1}} \tag{51}$$

This pattern is continued until the rest of the c's are all equal to zero. Then the d row is formed by using the s^{n-1} and s^{n-2} rows. These contants are

$$d_1 = \frac{c_1 b_{n-3} - b_{n-1}c_2}{c_1} \tag{52}$$

$$d_2 = \frac{c_1 b_{n-5} - b_{n-1}c_3}{c_1} \tag{53}$$

This is continued until no more d terms are present. The rest of the rows are formed in this way, down to the s^0 row. The complete array is triangular, ending with the s^0 row. Notice that the s^1 and s^0 rows contain only one term each. Once the array has been found, Routh's criterion states that the number of roots of the characteristic equation with position real parts is equal to the number of changes of sign of the coefficients in the first column. Therefore the system is stable if all terms in the first column have the same sign. Reference 17 shows that a system is unstable if there is a negative element in any position in any row.

The following example illustrates this criterion:

$$Q(s) = s^5 + s^4 + 10s^3 + 72s^2 + 152s + 240 \tag{54}$$

The Routhian array is formed by using the procedure described above:

$$\begin{array}{lccc} s^5| & 1 & 10 & 152 \\ s^4| & 1 & 72 & 240 \\ s^3| & -62 & -88 & \\ s^2| & 70.6 & 240 & \\ s^1| & 122.6 & & \\ s^0| & 240 & & \end{array} \tag{55}$$

In the first column there are two changes of sign, from 1 to -62 and from -62 to 70.6; therefore $Q(s)$ has two roots in the right-half s plane. Note that this criterion gives the number of roots with positive real parts but does not tell the values of the roots. If Eq. (54) is factored, the roots are $s = -3$, $s_{2,3} = -1 \pm j\sqrt{3}$, and $s_{4,5} = +2 \pm j4$. This confirms that there are two roots with positive real parts. The Routh criterion does not distinguish between real and complex roots.

Theorem 1. Division of a row: The coefficients of any row may be multiplied or divided by a positive number without changing the signs of the first column. The labor of evaluating the coefficients in Routh's array can therefore be reduced by multiplying or dividing any row by a constant. This may result, for example, in reducing the size of the coefficients and therefore simplifying the evaluation of the remaining coefficients.

Theorem 2. A zero coefficient in the first column: When the first term in a row is zero but not all the other terms in that row are zero, the following methods (14) can be used:

1. Substitute $s = 1/x$ in the original equation; then solve for the roots of x with positive real parts. The number of roots x with positive real parts will be the same as the number of s roots with positive real parts.
2. Multiply the original polynomial by the factor $(s + 1)$ which introduces an additional negative root. Then form the Routhian array for the new polynomial.

The first method is illustrated in the following example:

$$Q(s) = s^4 + s^3 + 2s^2 + 2s + 5 \tag{56}$$

The Routhian array is

$$\begin{array}{lccc} s^4| & 1 & 2 & 5 \\ s^3| & 1 & 2 & \\ s^2| & 0 & 5 & \end{array} \tag{57}$$

The zero in the first column prevents completion of the array. The following methods overcome this problem.

Method 1. Letting $s = 1/x$ and rearranging the polynomial gives

$$Q(x) = 5x^4 + 2x^3 + 2x^2 + x + 1 \tag{58}$$

The new Routhian array is

$$\begin{array}{lccc} s^4| & 5 & 2 & 1 \\ s^3| & 2 & 1 & \\ s^2| & -1 & 2 & \\ s^1| & 5 & & \\ s^0| & 2 & & \end{array} \tag{59}$$

There are two changes of sign; therefore there are two roots of x in the right-half s plane. The number of roots of s with positive real parts is also two. This method does not work when the coefficients of $Q(s)$ and of $Q(x)$ are identical.

Method 2

$$Q_1(s) = Q(s)(s+1) = s^5 + 2s^4 + 3s^3 + 4s^2 + 7s + 5 \tag{60}$$

The reader may obtain the Routhian array which has two changes of sign in the first column, so there are two zeros of $Q(s)$ with positive real parts. Thus, the same result is obtained by both methods. An additional method is described in Ref. 20.

Theorem 3. When all the coefficients of one entire row are zero, the procedure is as follows:

1. The auxiliary equation can be formed from the preceding row, as shown below.

2. The Routhian array can be completed by replacing the all-zero row with the coefficients obtained by differentiating the auxiliary equation.
3. The roots of the auxiliary equation are also roots of the original equation. These roots occur in pairs and are the negative of each other. Therefore, these roots may be imaginary (complex conjugates) or real (one positive and one negative), may lie in quadruplets (two pairs of complex-conjugate roots), and so on.

Consider the system which has the characteristic equation

$$q(s) = s^4 + 2s^3 + 11s^2 + 18s + 18 = 0 \tag{61}$$

$$\begin{array}{lccl} s^4| & 1 & 11 & 18 \\ s^3| & 2 & 18 & \\ \overline{s}^3| & 1 & 9 & \text{(After dividing the } s^3 \text{ row by 2)} \\ s^2| & 2 & 18 & \\ \overline{s}^2| & 1 & 9 & \text{(After dividing the } s^2 \text{ row by 2)} \\ s^1| & 0 & & \end{array} \tag{62}$$

The presence of a zero row for s^1 indicates that there are roots that are the negatives of each other. The next step is to form the auxiliary equation from the preceding row, which is the s^2 row. The highest power of s is s^2, and only even powers of s appear. Therefore the auxiliary equation is $s^2 + 9 = 0$. The roots of this equation are $s = \pm j3$. These are also roots of the original equation. The presence of imaginary roots indicates that the output includes a sinusoidally oscillating component. For a sinusoidal input with the frequency corresponding to the imaginary root, the response is unbounded. Thus, the system with imaginary roots is considered unstable.

To complete the Routhian array, the auxiliary equation is differentiated and is

$$2s + 0 = 0 \tag{63}$$

The coefficients of this equation are inserted in the s^1 row, and the array is then completed:

$$\begin{array}{ll} s^1| & 2 \\ s^0| & 9 \end{array} \tag{64}$$

Since there are no changes of sign in the first column, there are no roots with positive real parts.

In feedback systems, the ratio of the output to the input does not have an explicitly factored denominator (see the section entitled "Design Using the Root Locus"). An example of such a function is

$$\frac{Y(s)}{R(s)} = \frac{P(s)}{Q(s)} = \frac{K(s+2)}{s(s+5)(s^2+2s+5)+K(s+2)} \tag{65}$$

The value of K is an adjustable parameter in the system and may be positive or negative. The value of K determines the location of the poles and therefore the stability of the system. It is important to know the range of values of K for which the system is stable. This information must be obtained from the characteristic equation, which is

$$Q(s) = s^4 + 7s^3 + 15s^2 + (25+K)s + 2K \tag{66}$$

The coefficients must all be positive in order for the zeros of $Q(s)$ to lie in the left half of the s plane, but this is not a sufficient conditon for stability. The first column of the Routhian array permits evaluation of precise boundaries for K. It is left as an exercise for the reader to show that the closed-loop system is stable for $0 < K < 28.1$.

DESIGN USING THE ROOT LOCUS

As described in the section entitled "Location of Poles and Stability," the stability of linear time-invariant (LTI) systems is determined by the locations of the system poles. Any pole in the right-half plane leads to an exponentially increasing transient which therefore means that the system is unstable. An effective method for designing a feedback control is by use of the root locus. Figure 7 shows a block diagram of a single-input single-output feedback system expressed in terms of the Laplace transform variable s. $G(s)$ is the forward transfer function representing the dynamics of the "plant". Also, $H(s)$ is the feedback transfer function representing the dynamics in that path. The feedback signal $B(s)$ is compared with the command input signal $R(s)$, and their difference is called the actuating signal. The overall closed-loop transfer function is $C(s)/R(s)$, where $R(s)$ is the control system input and $C(s)$ is the controlled output signal.

For the root-locus method the open-loop transfer function can be put in the form

$$G(s)H(s) = \frac{K(s+a_1)\cdots(s+a_h)\cdots(s+a_w)}{s^m(s+b_1)(s+b_2)\cdots(s+b_c)\cdots(s+b_u)} \tag{67}$$

where a_h and b_c may be real or complex numbers and may lie in either the left-half or right-half s plane. The value of K may be either positive or negative. Equation (67) can be rewritten as

$$G(s)H(s) = \frac{K(s-z_1)\cdots(s-z_w)}{s^m(s-p_1)\cdots(s-p_u)} = \frac{K\prod_{h=1}^{w}(s-z_h)}{s^m\prod_{c=1}^{u}(s-p_c)} \tag{68}$$

where Π indicates a product of terms. The degree of the numerator is w and that of the denominator is $m + u = n$. The z_h are the zeros and the p_c are the poles of $G(s)H(s)$. In the form shown in Eq. (68), with the coefficients of s all equal to unity, the K is defined as the loop sensitivity.

The underlying principle of the root locus is that the poles of the control ratio $C(s)/R(s)$ are related to the zeros and poles of the open-loop transfer function $G(s)H(s)$ and to the loop

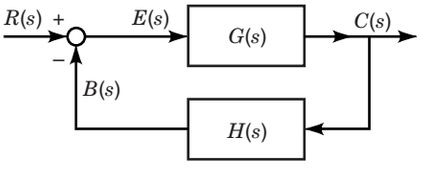

Figure 7. Block diagram of a feedback system.

sensitivity K. This is shown as follows. Let

$$G(s) = \frac{N_1(s)}{D_1(s)}, \qquad H(s) = \frac{N_2(s)}{D_2(s)} \tag{69}$$

Thus

$$\frac{C(s)}{R(s)} = M(s) = \frac{A(s)}{B(s)} = \frac{G(s)}{1 + G(s)H(s)} \tag{70}$$

where

$$B(s) \equiv 1 + G(s)H(s) = \frac{D_1D_2 + N_1N_2}{D_1D_2} \tag{71}$$

Rationalizing Eq. (70) gives

$$\frac{C(s)}{R(s)} = M(s) = \frac{N_1D_2}{D_1D_2 + N_1N_2} = \frac{P(s)}{Q(s)} \tag{72}$$

From Eqs. (71) and (72) it is seen that the zeros of $B(s)$ are equal to the poles of $M(s)$, and they determine the form of the closed-loop system response. The degree of the numerator of $B(s)$ is equal to $m + u$. Therefore $B(s)$ has $n = m + u$ finite zeros. The roots of $B(s) = 0$, which is the characteristic equation of the closed-loop system, must satisfy the equation

$$B(s) \equiv 1 + G(s)H(s) = 0 \tag{73}$$

These roots must satisfy the equation

$$G(s)H(s) = \frac{K(s - z_1)\cdots(s - z_w)}{s^m(s - p_1)\cdots(s - p_u)} = -1 \tag{74}$$

Thus, as the loop sensitivity K assumes values from zero to infinity, the open-loop transfer function must always be equal to -1. The corresponding values of s which satisfy Eq. (74) for any value of K are the poles of $M(s)$. The plots of these values of s are defined as the *root locus* of $M(s)$. They satisfy the following conditions:

For $K > 0$:

$$\text{Magnitude condition:} \quad |G(s)H(s)| = 1 \tag{75}$$

Angle condition:

$$\angle G(s)H(s) = (1 + 2h)180^\circ, \qquad h = 0, \pm1, \pm2, \ldots \tag{76}$$

For $K < 0$:

$$\text{Magnitude condition:} \quad |G(s)H(s)| = 1 \tag{77}$$

Angle condition:

$$\angle G(s)H(s) = h360^\circ, \qquad h = 0, \pm1, \pm2, \ldots \tag{78}$$

Thus, the root locus method provides a plot of the variation of each of the poles of $C(s)/R(s)$ in the complex s plane as the loop sensitivity K is varied from zero to infinity. All the angles are measured as positive in the counterclockwise sense. Since $G(s)H(s)$ usually has more poles than zeros, it is convenient to express the angle condition for the generalized form of $B(s)H(s)$ given in Eq. (67) as

$$|K| = \frac{|s^m| \cdot |s - p_1| \cdots |s - p_u|}{|s - z_1| \cdots |s - z_w|} \tag{79}$$

$$\begin{aligned} \beta &= \sum(\text{angles of denominator terms}) \\ &\quad - \sum(\text{angles of numerator terms}) \\ &= \begin{cases} (1 + 2h)180^\circ & \text{for } K > 0 \\ h360^\circ & \text{for } K < 0 \end{cases} \end{aligned} \tag{80}$$

Design Application: Root Locus Procedures

The procedure for obtaining the root locus is to first put the open-loop transfer function $G(s)H(s)$ into the form shown in Eq. (68). The poles and zeros are then plotted in the $s = \sigma + j\omega$ plane. Then the angle condition given in Eq. (80), for both $K > 0$ and $K < 0$, is used to obtain the complete root locus. Then the root locus can be calibrated in terms of the loop sensitivity K by using the magnitude condition given in Eq. (79). The dominant roots are those that contribute the most to the overshoot and the settling time of the closed-loop system response. Selection of the dominant complex roots on the root locus is based on the characteristics of a simple second-order system with a unit step input (8). The peak overshoot M_p is related to the damping ratio ζ by

$$M_p = 1 + \exp\left(-\frac{\zeta\pi}{\sqrt{1 - \zeta^2}}\right) \tag{81}$$

A specification based on M_p yields a required value of ζ. Roots with a specified value of ζ occur at the intersection of the root locus with a straight line drawn at an angle η drawn from the negative real axis, where

$$\eta = \cos^{-1}\zeta \tag{82}$$

An alternate specification is based on the desired settling time T_s. The real part of the desired complex root is given by

$$\sigma = \frac{\text{number of time constants}}{T_s} \tag{83}$$

The transient coefficient $e^{-\sigma t}$ decays to, and stays within a specified percentage of, the final value, in the time T_s. The latter is expressed as the number of time constants obtained from Eq. (83). For example, using 2% of the final value as the criterion requires four time constants for the selected root.

Root Locus Construction Rules

The following properties facilitate the drawing of the root locus:

Rule 1. The number of branches of the root locus is equal to the number of poles of the open-loop transfer function.

Rule 2. For positive values of K, the root locus exists on those portions of the real axis for which the total number of real poles and zeros to the right is an odd number. For negative values of K, the root locus exists on those portions of the real axis for which the total number of

real poles and zeros to the right is an even number (including zero).

Rule 3. The root locus starts ($K = 0$) at the open-loop poles and terminates ($K = \pm\infty$) at the open-loop zeros or at infinity.

Rule 4. The angles of the asymptotes of the root locus branches that end at infinity are determined by a

$$\gamma = \frac{(1+2h)180^\circ \text{ for } K>0 \qquad h360^\circ \text{ for } K<0}{\begin{matrix}[\text{Number of poles of} \\ G(s)H(s)]\end{matrix} \quad - \quad \begin{matrix}[\text{Number of zeros of} \\ G(s)H(s)]\end{matrix}} \tag{84}$$

Rule 5. The real-axis intercept of the asymptotes is

$$\sigma_o = \frac{\sum_{c=1}^{n} \text{Re}(p_c) - \sum_{h=1}^{w} \text{Re}(z_h)}{n-w} \tag{85}$$

Rule 6. The breakaway point for the locus between two poles on the real axis (or the break-in point for the locus between two zeros on the real axis) can be determined by taking the derivative of the loop sensitivity K with respect to s. Equate this derivative to zero and find the roots of the resulting equation. The root that occurs between the poles (or the zeros) is the breakaway (or break-in) point.

Rule 7. For $K > 0$ the angle of departure from a complex pole is equal to 180° minus the sum of the angles from the other poles plus the sum of the angles from the zeros. Any of these angles may be positive or negative. For $K < 0$ the departure angle is 180° from that obtained for $K > 0$. For $K > 0$ the angle of approach to a complex zero is equal to the sum of the angles from the poles minus the sum of the angles from the other zeros minus 180°. For $K < 0$ the approach angle is 180° from that obtained from $K > 0$.

Rule 8. The imaginary-axis crossing of the root locus can be determined by forming the Routhian array for the closed-loop characteristic equation. Equate the s^1 row to zero and form the auxiliary equation from the s^2 row. The roots of the auxiliary equation are the imaginary-axis crossover points.

Rule 9. The selection of the dominant roots of the characteristic equation is based on the specifications that give the required system performance; that is, it is possible to evaluate σ, ω_d, and ζ, which are used to select the location of the desired dominant roots. The loop sensitivity for these roots is determined by means of the magnitude condition. The remaining roots are then determined to satisfy the same magnitude condition.

Rule 10. For those open-loop transfer functions for which $w \leq n - 2$, the sum of the closed-loop roots is equal to the sum of the open-loop poles. Thus, once the dominant roots have been located,

$$\sum_{j=1}^{n} p_j = \sum_{j=1}^{n} r_j \tag{86}$$

can be used to find one real or two complex roots. Factoring known roots from the characteristic equation can also simplify the work of finding the remaining roots.

A root-locus digital-computer program (8,21) will produce an accurate calibrated root locus. This considerably simplifies the work required for the system design. By specifying ζ for the dominant roots or K_m, a computer program can determine all the roots of the characteristic equation.

Root Locus Example. Given here is the unity feedback system with

$$\begin{aligned} G(s) &= \frac{K_1}{s(s/25+1)(s^2/2600+s/26+1)} \\ &= \frac{65{,}000K_1}{s(s+25)(s^2+100s+2600)} \\ &= \frac{K}{s^4+125s^3+5{,}000s^2+65{,}000s} \end{aligned} \tag{87}$$

where $K = 65{,}000K_1$.

Specification: Find $C(s)/R(s)$ with $\zeta = 0.5$ for the dominant roots (roots closest to the imaginary axis).

1. The poles of $G(s)/H(s)$ are plotted on the s plane in Fig. 8. The values of the poles are: $s = 0, -25, -50 + j10, -50 - j10$. The system is completely unstable for $K < 0$. Therefore, this example is designed only for the condition $K > 0$.
2. There are four branches of the root locus. Since there are no zeros, all branches end at infinity for $K = \pm\infty$.
3. The locus exists on the real axis between 0 and -25.
4. The angles of the asymptotes are

$$\gamma = \frac{(1+2h)180^\circ}{4} = \pm 45^\circ, \pm 135^\circ$$

5. The real-axis intercept of the asymptotes is

$$\sigma_o = \frac{0-25-50-50}{4} = -31.25$$

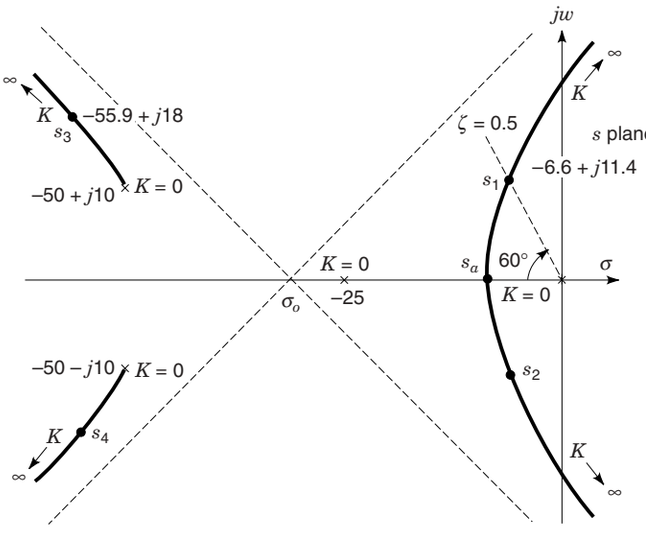

Figure 8. Root locus for

$$G(s)H(s) = \frac{65,000K_1}{s(s+25)(s^2+100s+2600)}$$

6. The breakaway point s_a on the real axis between 0 and -25 is found from the derivative $dK/ds = 0$, using Eq. (79) to obtain $K(s)$. The result is

$$-\frac{dK}{ds} = 4s^3 + 375s^2 + 10{,}200s + 65{,}000 = 0$$

This yields $s_a = -9.15$.

7. The angle of departure from the pole $-50 + j10$ is obtained from

$$\phi_0 + \phi_1 + \phi_2 + \phi_{3D} = (1 + 2h)180^\circ$$

where the angles are measured from each open-loop pole to the pole $-50 + j10$. This yields $\phi_{3D} = 123.1^\circ$.

8. Using Eq. (72), the closed-loop transfer function is

$$\frac{C(s)}{R(s)} = \frac{65{,}000K_1}{s^4 + 125s^3 + 5{,}100s^2 + 65{,}000s + 65{,}000K_1} \quad (88)$$

The Routhian array for the denominator of $C(s)/R(s)$, which is the characteristic polynomial, is

s^4|	1	5100	$65{,}000K_1$
s^3|	1	520	(After division by 125)
s^2|	1	$14.2K_1$	(After division by 4580)
s^1|	$520 - 14.2K_1$		
s^0|	$14.2K_1$		

Pure imaginary roots exist when the s^1 row is zero. This occurs when $K_1 = 520/14.2 = 36.6$. The auxiliary equation is $s^2 + 14.2K_1 = 0$. The imaginary roots are $s = \pm j\sqrt{14.2K_1} = \pm j22.8$.

9. Additional points on the root locus are found by locating points that satisfy the angle condition

$$\angle s + \angle(s + 25) + \angle(s + 50 - j10) + \angle(s + 50 + j10) = (1 + 2m)180^\circ$$

The root locus is shown in Fig. 8.

10. The radial line for $\zeta = 0.5$ is drawn on the graph of Fig. 8 at the angle $\eta = \cos^{-1}\zeta = \cos^{-1}0.5 = 60^\circ$. The dominant complex roots obtained from the graph are

$$s_{1,2} = -6.6 \pm j11.4$$

11. Applying the magnitude condition of Eq. (79) yields

$$K = 65{,}000K_1 = |s| \cdot |s + 25| \cdot |s + 50 - j10| \cdot |s + 50 + j10|$$

For $s = -6.6 + j11.4$, $K = 598{,}800$ and $K_1 = 9.25$.

12. The other roots are evaluated to satisfy the magnitude conditon $K = 598{,}800$. The other roots of the characteristic equation are

$$s = -55.9 \pm j18.0$$

13. The overall closed-loop transfer function is

$$\frac{C(s)}{R(s)} = \frac{598{,}800}{(s + 6.6 \pm j11.4)(s + 55.9 \pm j18)} \quad (89)$$

14. With a unit step input, $R(s) = 1/s$, the output response can be obtained from a CAD program (8,21):

$$c(t) = 1 + 1.4e^{-6.6t}\sin(11.4t - 143.8^\circ) + 0.2e^{-55.9t}\sin(18.0t - 123.7^\circ) \quad (90)$$

The figures of merit which describe the significant step response characteristics are $t_s = 0.64$ s, $t_p = 0.31$ s, $M_p = 1.56$, and $c(t)_{ss} = 1.00$.

The root locus design method ensures more than system stability. It also permits selection of the closed-loop system poles that best satisfy the desired performance specifications.

FREQUENCY RESPONSE

The objective in using feedback is threefold. The first purpose is to achieve stability and/or a specified degree of stability margin. The second goal is to reduce the sensitivity to parameter variation. The third goal is disturbance rejection. The previous section introduces the root locus method of design. The closed-loop poles are assigned in the left-half plane so that the performance is stable, and they are located on the root locus in positions that best achieve the desired performance criteria. This section introduces the frequency response methods that ensure closed-loop system stability by application of the Nyquist criterion and meet desired figures of merit. The gain of the forward transfer function is the adjustable parameter which is used in the design. The feedback control system is shown in Fig. 7. For a linear time-invariant (LTI) system, the open-loop transfer function in the s domain is shown in Eq. (68), and the closed-loop transfer function is represented by Eqs. (70) and (72).

The forward frequency transfer function is written in the generalized form:

$$G(j\omega) = \frac{K_m(1 + j\omega T_1)(1 + j\omega T_2)\cdots(1 + j\omega T_w)}{(j\omega)^m(1 + j\omega T_a)[1 + (2\zeta/\omega_n)j\omega + (1/\omega_n^2)(j\omega)^2]\cdots} = K_m G'(j\omega) \quad (91)$$

where m defines the system type, K_m is the gain constant, and $G'(j\omega)$ is the forward transfer function with unity gain. There are two forms typically used for the plot of $G(j\omega)$. In the first category is the Nyquist plot which is the output–input ratio in polar coordinates. The Nyquist stability criterion is applied to this plot to determine the closed-loop system stability. The second category is the pair of Bode plots. One plot is the magnitude in decibels versus the logarithm of frequency, and the second plot is the angle versus the logarithm of frequency. Typically, semilog graph paper is used so that it is not necessary to obtain the logarithm of frequency. The data from the Bode plot is then plotted on a *Nichols* plot in rectangular coordinates: log magnitude versus angle. The Nichols chart is a set of contours which are drawn on the Nichols plot and is used to both ensure stability and to adjust the gain in order to meet the closed-loop specifications. A digital computer CAD program (8,21) provides considerable assistance in obtaining these plots and performing the design.

Direct Polar-Plot Characteristics

To obtain the direct polar plot of a system's forward transfer function, the following characteristics are used to determine the key parts of the curve.

Step 1. The forward transfer function has the general form shown in Eq. (91). For this transfer function the system type is equal to the value of m and determines the portion of the polar plot representing the $\lim_{\omega\to 0} G(j\omega) = 0\angle(w - m - u)90°$. The low-frequency polar-plot characteristics (as $\omega \to 0$) of the different system types are summarized in Fig. 9. The angle at $\omega = 0$ is $m(-90°)$. The arrow on the polar plots indicates the direction of increasing frequency.

Step 2. The high-frequency end of the polar plot can be determined as follows:

$$\lim_{\omega\to+\infty} G(j\omega) = 0\angle(w - m - u)90° \tag{92}$$

Note that since the degree of the denominator of Eq. (91) is usually greater than the degree of the numerator, the high-frequency point ($\omega = \infty$) is approached (i.e., the angular condition) in the clockwise sense. The plot ends at the origin, tangent to the axis determined by Eq. (92). Tangency may occur on either side of the axis.

Step 3. The asymptote that the low-frequency end approaches, for a Type 1 system, is determined by taking the limit as $\omega \to 0$ of the real part of the transfer function.

Step 4. The frequencies at the points of intersection of the polar plot with the negative real axis and the imaginary axis are determined, respectively, by setting

$$\begin{aligned} \mathrm{Im}[G(j\omega)] &= 0 \\ \mathrm{Re}[G(j\omega)] &= 0 \end{aligned} \tag{93}$$

Step 5. If there are no frequency-dependent terms in the numerator of the transfer function, the curve is a smooth one in which the angle of $\mathbf{G}(j\omega)$ continuously decreases as ω goes from 0 to ∞. With time constants in the numerator, and depending upon their values, the angle may not continuously vary in the same direction, thus creating "dents" in the polar plot.

Step 6. As is seen later in this chapter, it is important to know the exact shape of the polar plot of $\mathbf{G}(j\omega)$ in the vicinity of the $-1 + j0$ point and the crossing point on the negative real axis.

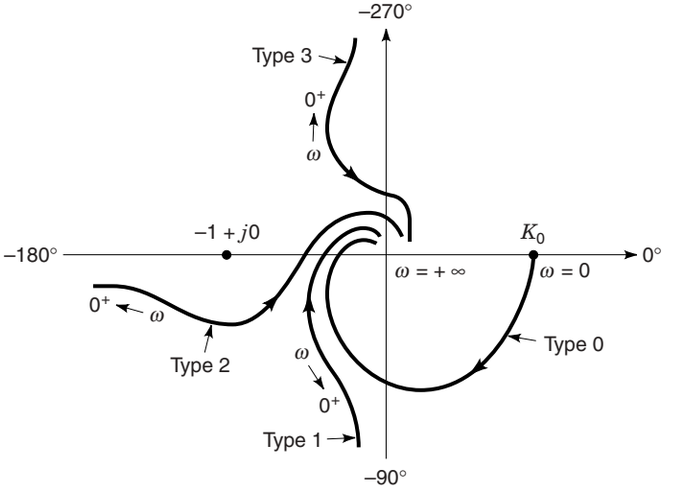

Figure 9. A summary of direct polar plots of different types of systems.

Nyquist's Stability Criterion

A system designer must be sure that the closed-loop system he designs is stable. The Nyquist stability criterion (22,23) provides a simple graphical procedure for determining closed-loop stability from the frequency–response curves of the open-loop transfer function $\mathbf{G}(j\omega)\mathbf{H}(j\omega)$. The application of this method in terms of the polar plot is covered in this section; application in term of the log magnitude–angle (Nichols) diagram is covered in a later section. The closed-loop transfer function of the system is given in Eqs. (70) and (72). The characteristic equation formed from the denominator of this closed-loop transfer function is

$$B(s) = 1 + G(s)H(s) = \frac{D_1D_2 + N_1N_2}{D_1D_2} = 0 \tag{94}$$

For a stable system the roots of the characteristic equation must not lie in the right-half s plane or on the $j\omega$ axis. Note that the numerator and denominator of $B(s)$ have the same degree. The poles of the open-loop transfer function $G(s)H(s)$ are the poles of $B(s)$. Since the denominator of $C(s)/R(s)$ is the same as the numerator of $B(s)$, the condition for stability may therefore be restated as: For a stable system, none of the zeros of $B(s)$ can lie in the right-half s plane or on the imaginary axis. Nyquist's stability criterion relates the number of zeros and poles of $B(s)$ that lie in the right-half s plane to the polar plot of $G(s)H(s)$.

In this analysis it is assumed that the control system's range of operation is confined to the linear region. This yields a set of linear differential equations which describe the dynamic performance of the systems. Because of the physical nature of feedback control systems, the order of the denominator D_1D_2 is equal to or greater than the order of the numerator N_1N_2 of the open-loop transfer function $G(s)H(s)$. This means that $\lim_{s\to\infty} G(s)H(s) \to 0$ or a constant.

A rigorous mathematical derivation of Nyquist's stability criterion is based on complex variable theory. A qualitative approach to Nyquist's stability criterion is presented for the special case that $B(s)$ is a rational fraction. The characteristic function $B(s)$ given by Eq. (94) can be rationalized, factored, and then written in the form

$$B(s) = \frac{(s - Z_1)(s - Z_2)\cdots(s - Z_n)}{(s - p_1)(s - p_2)\cdots(s - p_n)} = \frac{P(s)}{Q(s)} \tag{95}$$

where $Z_1, Z_2, \ldots, Z_n$ are the zeros and $p_1, p_2, \ldots, p_n$ are the poles. The poles p_i are the same as the poles of the open-loop transfer function $G(s)H(s)$ and include the s term for which $p = 0$, if it is present.

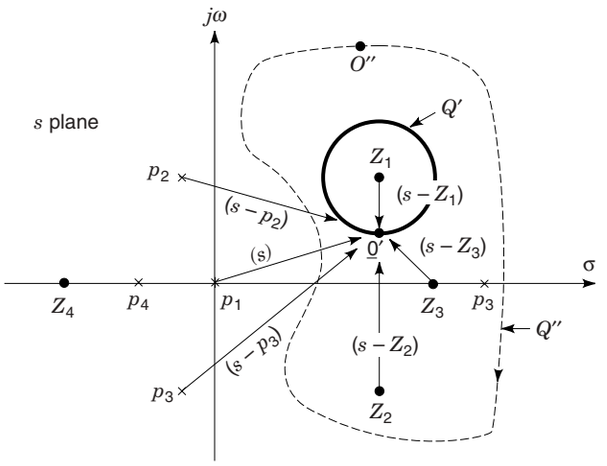

Figure 10. A plot of some poles and zeros of Eq. (95).

In Fig. 10 some poles and zeros of a generalized function $B(s)$ are drawn on the s plane. Also, an arbitrary closed contour Q' is drawn which encloses the zero Z_1. To the point O' on Q', whose coordinates are $s = \sigma + j\omega$, are drawn directed line segments from all the poles and zeros. The lengths of these directed line segments are given by $s - Z_1$, $s - Z_2$, $s - p_1$, $s - p_2$, and so on. Not all the directed segments from the poles and zeros are indicated in the figure, because they are not necessary to proceed with this development. As the point O' is rotated clockwise once around the closed contour Q', the directed segment $s - Z_1$ rotates through a net clockwise angle of 360°. All the other directed segments rotate through a net angle of 0°. Thus, referring to Eq. (95), it is seen that the clockwise rotation of 360° for $s - Z_1$ is simultaneously realized by the function $B(s)$ for the enclosure of the zero Z_1 by the path Q'.

Consider now a larger closed contour Q'' which includes the zeros Z_1, Z_2, and Z_3 and the pole p_5. As a point O'' is rotated clockwise once around the closed curve Q'', each of the directed line segments from the enclosed pole and zeros rotates through a net clockwise angle of 360°. Since the angular rotation of the pole is experienced by the characteristic function $B(s)$ in its denominator, the net angular rotation realized by Eq. (95) must be equal to the net angular rotations due to the pole p_5 minus the net angular rotations due to the zeros Z_1, Z_2, and Z_3. Therefore, for this case, the net number of rotations N experienced by $B(s) = 1 + G(s)H(s)$ for the clockwise movement of point O'' once about the closed contour Q'' is equal to -2; that is,

$$N = (\text{number of poles enclosed}) - (\text{number of zeros enclosed}) = 1 - 3 = -2$$

where the minus sign denotes clockwise (cw) rotation. Also, for any closed path that may be chosen, all the poles and zeros that lie outside the closed path each contribute a net angular rotation of 0° to $B(s)$ as a point is moved once around this contour.

Generalizing Nyquist's Stability Criterion

Consider now a closed contour Q which encloses the whole right-half s plane (see Fig. 11), thus enclosing all zeros and poles of $B(s)$ that have positive real parts. As a consequence of the theory of complex variables used in the formal derivation, the contour Q must not pass through any poles or zeros of $B(s)$. When the results of the preceding discussion are applied to the contour Q, the following properties are noted:

1. The total number of clockwise rotations of $B(s)$ due to its zeros is equal to the total number of zeros Z_R in the right-half s plane.
2. The total number of counterclockwise rotations of $B(s)$ due to its poles is equal to the total number of poles P_R in the right-half s plane.
3. The net number of rotations N of $B(s) = 1 + G(s)H(s)$ about the origin is equal to its total number of poles P_R minus its total number of zeros Z_R in the right-half s plane. N may be positive (ccw), negative (cw), or zero.

The essence of these three properties can be represented by the equation

$$N = \text{Change in phase of } [1 + G(s)H(s)]/2\pi = P_R - Z_R \quad (96)$$

where counterclockwise rotation is defined as being positive and clockwise rotation is negative. In order for the characteristic function $B(s)$ to realize a net rotation N, the directed line segment representing $B(s)$ (see Fig. 12) must rotate about the origin $360N$ degrees, or N complete revolutions. Solving Eq. (96) for Z_R yields

$$Z_R = P_R - N \quad (97)$$

For a stable system, $B(s)$ can have no zeros Z_R in the right-half s plane. It is therefore concluded that, for a stable system, the net number of rotations of $B(s)$ about the origin must be counterclockwise and equal to the number of poles P_R that lie in the right-half plane. In other words, if $B(s)$ experiences a net clockwise rotation (i.e., if N is negative), this indicates that $Z_R > P_R$, where $P_R \geq 0$, and thus the closed-loop system is unstable. If there are zero net rotations, then $Z_R = P_R$ and the system may or may not be stable, according to whether $P_R = 0$ or $P_R > 0$.

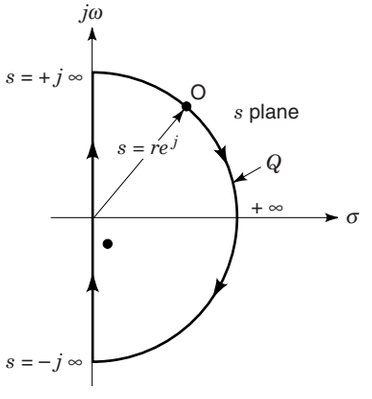

Figure 11. The contour that encloses the entire right-half s plane.

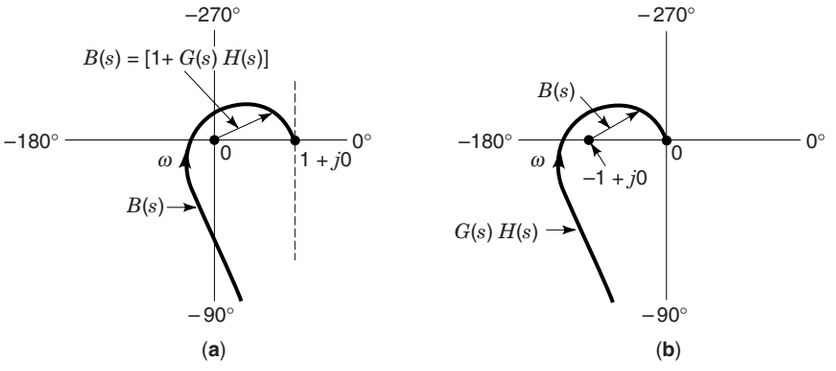

Figure 12. A change of reference for $B(s)$.

Obtaining a Plot of $B(s)$. Figures 12(a) and 12(b) show a plot of $B(s)$ and a plot of $G(s)H(s)$, respectively. By moving the origin of Fig. 12(b) to the $-1 + j0$ point, the curve is now equal to $1 + G(s)H(s)$, which is $B(s)$. Since $G(s)H(s)$ is known, this function is easily plotted, and then the origin is moved to the -1 point to obtain $B(s)$. In general, the open-loop transfer functions of many physical systems do not have any poles P_R in the right-half s plane. In this case, $Z_R = N$. Thus, for a stable system the net number of rotations about the $-1 + j0$ point must be zero when there are no poles of $G(s)H(s)$ in the right-half s plane. If the function $G(s)H(s)$ has some poles in the right-half s plane and the denominator is not in factored form, then the number P_R can be determined by applying Routh's criteiron to D_1D_2. The Routhian array gives the number of roots in the right-half s plane by the number of sign changes in the first column.

Analysis of Path Q. Nyquist's criterion requires that the whole right-half s plane must be encircled to ensure the inclusion of all poles or zeros in this portion of the plane. In Fig. 11 the entire right-half s plane is included by the closed path Q which is composed of the following two segments:

1. One segment is the imaginary axis from $-j\infty$ to $+j\infty$.
2. The other segment is a semicircle of infinite radius that encircles the entire right-half s plane.

The portion of the path along the imaginary axis is represented mathematically by $s = j\omega$. Thus, replacing s by $j\omega$ in Eq. (95) and letting ω take on all values from $-\infty$ to $+\infty$ gives the portion of the $B(s)$ plot corresponding to that portion of the closed contour Q which lies on the imaginary axis.

A requirement of the Nyquist criterion is that $\lim_{s\to\infty} G(s)H(s) \to 0$ or a constant. Therefore, as the point O moves along the segment of the closed contour represented by the semicircle of infinite radius, the corresponding portion of the $B(s)$ plot is a fixed point. As a result, the movement of point O along only the imaginary axis from $-j\infty$ to $+j\infty$ results in the same net rotation of $B(s)$ as if the whole contour Q were considered. In other words, all the rotation of $B(s)$ occurs while the point O goes from $-j\infty$ to $+j\infty$ along the imaginary axis.

Effect of Poles at the Origin on the Rotation of $B(s)$. Some transfer functions $G(s)H(s)$ have an s^m in the denominator. Since no poles or zeros can lie on the contour Q, the contour shown in Fig. 11 must be modified to the contour shown in Fig. 13(a). Consider the transfer function

$$G(s)H(s) = \frac{K_1}{s(1+T_1s)(1+T_2s)} \tag{98}$$

The point O is first moved along the negative imaginary axis from $s = -j\infty$ to a point where $s = -j\omega = 0^- \angle -90°$ becomes very small; that is, $s = -j\epsilon$. Then the point O moves along a semicircular path of radius $s = \epsilon e^{j\theta}$ in the right-half s plane, with a very small radius ϵ until it reaches the positive imaginary axis at $s = +j\omega = j0^+ = 0^+ \angle 90°$. From here the point O proceeds along the positive imaginary axis to $s = +j\infty$. Letting the radius approach zero, $\epsilon \to 0$, for the semicircle around the origin, ensures the inclusion of all poles and zeros in the right-half s plane. To complete the plot of $B(s)$, the effect of moving point O on this semicircle around the origin must be investigated.

For the semicircular portion of the path Q represented by $s = \epsilon e^{j\theta}$, where $\epsilon \to 0$ and $-90° < \theta < +90°$, Eq. (98) becomes

$$G(s)H(s) = \frac{K_1}{s} = \frac{K_1}{\epsilon e^{j\theta}} = \frac{K_1}{\epsilon} e^{-j\theta} = \frac{K_1}{\epsilon} e^{j\psi} \tag{99}$$

where $K_1/\epsilon \to \infty$ as $\epsilon \to 0$, and $\psi = -\theta$ goes from 90° to −90° as the directed line segment s goes counterclockwise from $\epsilon \angle -90°$ to $\epsilon \angle +90°$. Thus, in Fig. 13(b), the points $G(s)H(s)$ for $\omega \to 0^-$ and $\omega \to 0^+$ are joined by a semicircle of infinite radius in the first and fourth quadrants. Figure 13(b) shows the completed contour of $G(s)H(s)$ as the point O moves along the

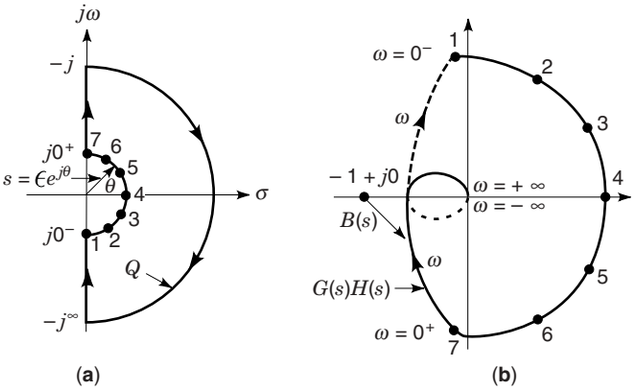

Figure 13. (a) The contour Q which encircles the right-half s plane. (b) Complete plot for Eq. (98).

modified contour Q of Fig. 13(a) in the s plane in the clockwise direction. When the origin is moved to the $-1 + j0$ point, the curve becomes $B(s)$. The plot of $B(s)$ in Fig. 13(b) does not encircle the $-1 + j0$ point; therefore $N = 0$. Also, from Eq. (98), $P_R = 0$. Thus, using Eq. (97), $Z_R = 0$ and the system is stable. If the gain K_1 is increased so that the intersection of $G(s)H(s)$ with the negative real axis occurs to the left of the -1 point, then $N = -2$ and $Z = 2$, with the result that the closed-loop system is unstable.

Nyquist Stability: Type 2 System. Transfer functions that have the term s^m in the denominator have the general form, as $\epsilon \to 0$,

$$G(s)H(s) = \frac{K_m}{s^m} = \frac{K_m}{\epsilon^m e^{jm\theta}} = \frac{K_m}{\epsilon^m} e^{-jm\theta} = \frac{K_m}{\epsilon^m} e^{jm\psi} \qquad (100)$$

where $m = 1, 2, 3, 4, \ldots$. With the reasoning used in the preceding example, it is seen from Eq. (100) that, as s moves from 0^- to 0^+, the plot of $G(s)H(s)$ traces m clockwise semicircles of infinite radius about the origin. For example, if $m = 2$, then, as θ goes from $-\pi/2$ to $+\pi/2$ in the s plane with radius, ϵ, $G(s)H(s)$ experiences a net rotation of $(2)(180°) = 360°$. Since the polar plots are symmetrical about the real axis, it is only necessary to determine the shape of the plot of $G(s)H(s)$ for a range of values of $0 < \omega < +\infty$. The net rotation N of the plot for the range of $-\infty < \omega < +\infty$ is twice that of the plot for the range of $0 < \omega < +\infty$. Consider the system

$$G(s)H(s) = \frac{K_2(1 + T_4 s)}{s^2(1 + T_1 s)(1 + T_2 s)(1 + T_3 s)} \qquad (101)$$

where $T_4 > T_1 + T_2 + T_3$. Figure 14 shows the mapping of $G(s)H(s)$ for the contour Q of the s plane. The word *mapping,* as used here, means that for a given point in the s plane there corresponds a given value of $G(s)H(s)$ or $B(s)$. The presence of the s^2 term in the denominator of Eq. (101) results in a net rotation of 360° in the vicinity of $s = 0$, as shown in Fig. 14. For the complete range of frequencies the net rotation is zero; thus, since $P_R = 0$, the system is stable. The value of N can be determined by drawing the line radiating from the $-1 + j0$ point (see Fig. 14) and noting one cw and one ccw crossing; thus, the sum of these crossings is $N = 0$. Like the previous example, this system can be made unstable by increasing the gain sufficiently for the $G(s)H(s)$ plot to cross the negative real axis to the left of the $-1 + j0$ point.

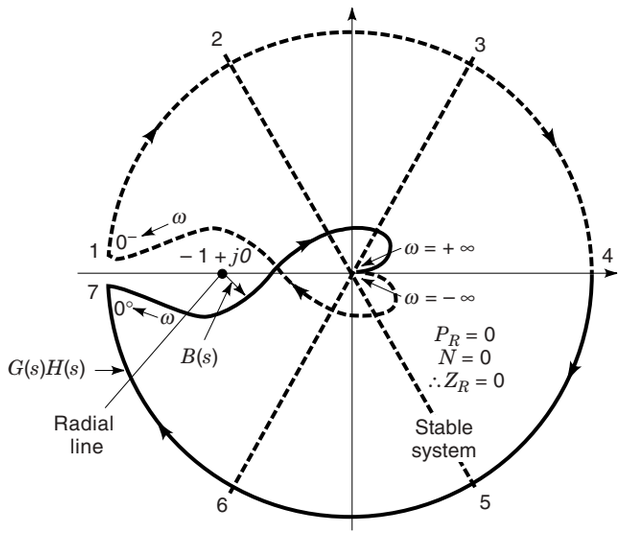

Figure 14. The complete polar plot for Eq. (101).

When $G(j\omega)H(j\omega)$ Passes Through the $-1 + j0$ Point. When the curve of $G(j\omega)H(j\omega)$ passes through the $-1 + j0$ point, the number of encirclements N is indeterminate. This corresponds to the condition where $B(s)$ has zeros on the imaginary axis. A necessary condition for applying the Nyquist criterion is that the path encircling the specified area must not pass through any poles or zeros of $B(s)$. When this condition is violated, the value for N becomes indeterminate and the Nyquist stability criterion cannot be applied. Simple imaginary zeros of $B(s)$ mean that the closed-loop system will have a continuous steady-state simusoidal component in its output which is independent of the form of the input. In addition, a sinusoidal input with the frequency equal to that of the imaginary zero produces an unbounded output. Therefore, this condition is considered unstable.

Nichols Plot and Stability

The *log magnitude* of the open-loop transfer function $\mathbf{G}(j\omega)\mathbf{H}(j\omega)$, abbreviated as Lm, is defined as 20 times the logarithm to the base 10:

$$\text{Lm } \mathbf{G}(j\omega) = 20 \log_{10} |\mathbf{G}(j\omega)| \quad \text{dB} \qquad (102)$$

where the Lm has the units of decibels (dB).

The *Nichols plot* is defined for the open-loop transfer function $\mathbf{G}(j\omega)\mathbf{H}(j\omega)$ as having the Lm on the vertical axis and the angle on the horizontal axis. This is an alternate to the polar plot and can be used to determine closed-loop system stability by use of the Nyquist criterion. The corresponding polar and Nichols plots for the transfer function of Eq. (98) are shown in Figs. 15(a) and 15(b). The data for plotting these curves can be obtained by use of a computer-aided design (CAD) program (8,21).

The log-magnitude–angle diagram is drawn by picking for each frequency the values of log magnitude and angle. The resultant curve has frequency as a parameter. The curve for the Eq. (98) is drawn in Fig. 15(b). Note that the point $-1 + j0$ on the polar plot becomes a series of points on the Nichols plot, having the values (0 dB, $k\pi$), where k takes on all odd integer values. Changing the gain raises or lowers the Lm $G(j\omega)$ curve without changing the angle characteristics. Increasing the gain raises the curve, and analysis of the polar plot has shown that this reduces stability.

The log-magnitude–angle diagram for $G(s)H(s)$ can be drawn for all values of s on the contour Q of Fig. 13(a). The resultant curve for minimum-phase systems is a closed contour. Nyquist's criterion can be applied to this contour by determining the number of points having the values 0 dB and odd multiples of 180°, which are enclosed by the curve of $G(s)H(s)$. This number is the value of N which is used in the equation $Z_R = N - P_R$ to determine the value of Z_R. As an example, consider a control system whose transfer function is given by Eq. (98). Its log-magnitude–angle diagram, for the contour Q, is shown in Fig. 16. From this figure it is seen that the value of N is zero. Thus $Z_R = N - P_R = 0$, and the system

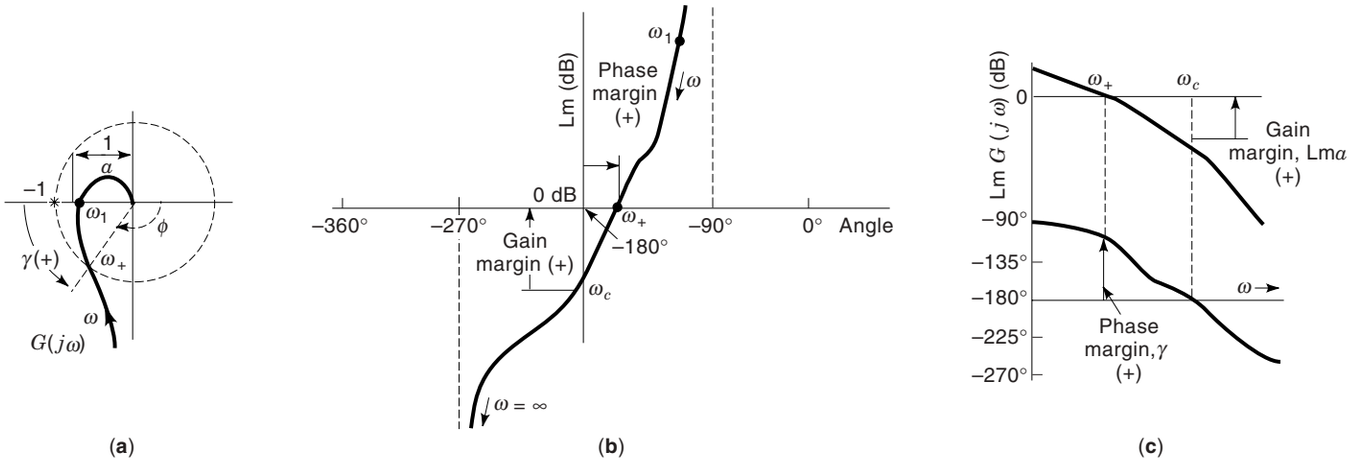

Figure 15. Frequency response plots: (a) Polar plot, (b) Nichols plot, and (c) Bode plots of $G(j\omega)$.

is stable. The log-magnitude–angle contour for a non-minimum-phase system does not close; thus it is difficult to determine the value of N. For these cases the polar plot is easier to use to determine stability.

It is not necessary for minimum-phase systems to obtain the complete log-magnitude–angle contour to determine stability. Only that portion of the contour is drawn representing $\mathbf{G}(j\omega)\mathbf{H}(j\omega)$ for the range of values $0^+ < \omega < \infty$. The stability is then determined from the position of the curve of $\mathbf{G}(j\omega)\mathbf{H}(j\omega)$ relative to the (0 dB, $-180°$) point. In other words, the curve is traced in the direction of increasing frequency—that is, walking along the curve in the direction of increasing frequency. The system is stable if the (0 dB, $-180°$) point is to the right of the curve. This is a simplified rule of thumb which is based on Nyquist's stability criterion for a minimum-phase system.

A conditionally stable system is one in which the curve crosses the $-180°$ axis at more than one point. Figure 17 shows the transfer-function plot for such a system with two stable and two unstable regions. The gain determines whether the system is stable or unstable.

Gain Margin and Phase Margin Stability from the Nichols Plot

The absolute stability of an LTI closed-loop system can be determined by applying the Nyquist stability theorem, using the representation of the open-loop transfer function $\mathbf{G}(j\omega)$ as a polar plot or as Nichols plot (log magnitude versus angle). Some measures of degree of stability can be expressed in terms of gain margin and phase margin. The following quantities are used to express these stability measures:

Gain Crossover. This is the point on the plot of the transfer function at which the magnitude of $\mathbf{G}(j\omega)$ is unity [Lm $\mathbf{G}(j\omega) = 0$ dB]. The frequency at gain crossover is called the phase-margin frequency ω_ϕ.

Phase Margin. This is 180° plus the negative trigonometrically considered angle of the transfer function at the gain-crossover point. It is designated as the angle γ, which can be expressed as $\gamma = 180° + \phi$, where $\angle\mathbf{G}(j\omega_\phi) = \phi$ is negative.

Phase Crossover. This is the point on the plot of the transfer function at which the phase angle is $-180°$. The fre-

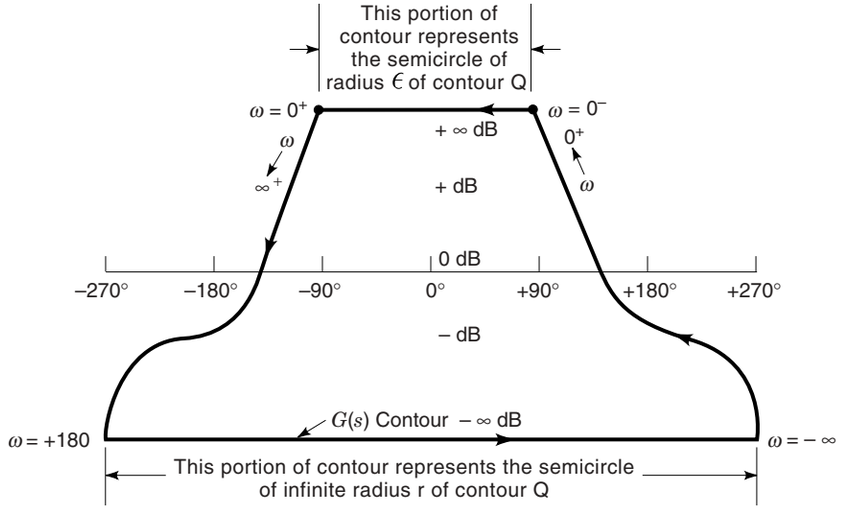

Figure 16. The log-magnitude–angle contour for the minimum-phase system of Eq. (98).

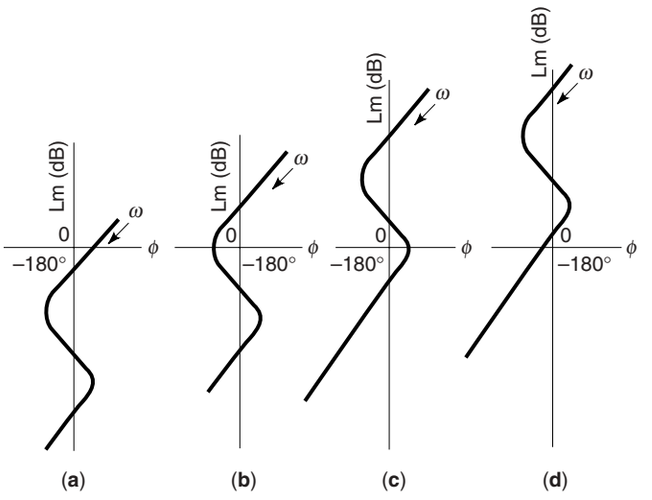

Figure 17. Log-magnitude–angle diagram for a conditionally stable system: (a, c) stable; (b, d) unstable.

quency at which phase crossover occurs is called the gain-margin frequency ω_c.

Gain Margin. The gain margin is the factor a by which the gain must be changed in order to produce instability. Expressed in terms of the transfer function at the frequency ω_c, it is

$$|\mathbf{G}(\omega_c)|a = 1 \tag{103}$$

On the polar plot of $\mathbf{G}(j\omega)$, the value at ω_c, as shown in Fig. 15(a), is $|\mathbf{G}(j\omega_c)| = 1/a$. In terms of the Lm in dB, the gain margin is

$$\text{Lm } a = -\text{Lm } \mathbf{G}(j\omega_c) \tag{104}$$

This is shown in Fig. 15(b) as the amount by which the plot of $\mathbf{G}(j\omega)$ must be raised so that it goes through the 0 dB, 180° $(-1 + j0)$ point.

These quantities are illustrated in Fig. 15 on both the polar and the Nichols plots. Note the algebraic sign associated with these quantities as marked on the curves. Both cases shown on Figs. 15(a) and 15(b) represent stable systems.

The phase margin angle is the amount of phase shift at the frequency ω_ϕ that would just produce instability. The phase margin for minimum-phase systems must be positive for a stable system, whereas a negative phase margin means that the system is unstable.

It can be shown (8) that the phase margin angle γ is related to the effective damping ratio ζ of the system. Satisfactory response is usually obtained with a phase margin of 45° to 60°. As an individual gains experience and develops his own particular technique, the desirable value of γ to be used for a particular system becomes more evident. This guideline for system performance applies only to those systems where closed-loop behavior is that of an equivalent second-order system. The gain margin must be positive when expressed in decibels (greater than unity as a numeric) for a stable system. A negative gain margin means that the system is unstable. The damping ratio ζ of the closed-loop system is also related to the gain margin. However, the phase margin gives a better estimate of damping ratio, and therefore of the transient overshoot of the closed-loop system, than the gain margin.

The phase-margin frequency ω_ϕ, phase margin γ, crossover frequency ω_c, and the gain margin Lm a are readily identified on the Nichols plot, as shown in Fig. 15(b).

Bode Plots (Logarithmic Plots)

The Bode plots consist of two components. One plot is for the log magnitude [Lm $G(j\omega) = 20 \log \mathbf{G}(j\omega)$] versus log ω, and the second plot is for the angle of $\mathbf{G}(j\omega)$ versus log ω. The Bode plots are often used to represent the open-loop transfer function. The log magnitude of $\mathbf{G}(j\omega)$ converts the operations of multiplication and divisoin to addition and subtraction, respectively. Each factor of a transfer function [see Eq. (91)], plotted versus log ω, has a distinctive characteristic which is readily recognized.

Constant K_m. The Lm K_m is a constant which is independent of frequency. Thus, it is a horizontal line and serves to raise or lower Lm $\mathbf{G}(j\omega)$, when it is larger than or smaller than unity, respectively.

$j\omega$ Factor. The Lm $\mathbf{G}(j\omega)$ has a positive slope of 20 dB per decade when this factor appears in the numerator. Its slope is negative when it appears in the denominator of $\mathbf{G}(j\omega)$. The transfer function *type* is related to the low frequency slope of the Lm plot. A zero slope indicates a Type 0 system, a slope of −20 dB/decade indicates a Type 1 system, and so on.

$1 + j\omega T$ Factor. The Lm$(1 + j\omega T)$ has a *corner frequency* at $\omega_{cf} = 1/T$. At frequencies below ω_{cf} the asymptote is the 0 dB line. Above ω_{cf} the asymptote is a straight line that passes through $\omega = 1/T$ at zero dB, with a slope of 20 dB/decade. When $(1 + j\omega T)$ appears in the numerator of $\mathbf{G}(j\omega)$, the slope of the asymptote is positive. When it appears in the denominator, the slope of the asymptote is negative.

Quadratic Factor $[1 + 2\zeta j\omega/\omega_n + (j\omega/\omega_n)^2]$. The log magnitude of the quadratic factor has a corner frequency at $\omega_{cf} = \omega_n$. At frequencies below ω_{cf} the asymptote is the 0 dB line. Above ω_{cf} the asymptote is a straight line that passes through ω_{cf} at 0 dB with a slope of 40 dB/decade. When the quadratic factor is in the denominator, the slope of the asymptote is negative. When it appears in the numerator, the slope of the asymptote is positive. Because this quadratic factor has the damping ratio ζ as an additional variable, there is a family of log magnitude plots which depend on the value of ζ.

The gain margin and the phase margin angle can be determined from the Bode plots, as shown in Fig. 15(c). Adjusting the gain results in raising or lowering the log magnitude curve, without changing the angle curve of $\mathbf{G}(j\omega)$. This permits the designer to change the phase margin frequency ω_ϕ, the phase margin angle γ, and the gain margin. For a stable closed-loop system, it is necessary for both the phase margin angle and the gain margin to be positive.

Experimental Determination of Transfer Function

The log-magnitude–phase-angle diagram is of great value when the mathematical expression for the transfer function

of a given system is not known. The magnitude and angle of the ratio of the output to the input can be obtained experimentally for a steady-state sinusoidal input signal at a number of frequencies. These data values are used to obtain the exact log-magnitude–angle diagram. Asymptotes are drawn on the exact log magnitude curve, using the fact that their slopes must be multiples of 20 dB/decade. From these asymptotes, the system type and the approximate time constants are determined. Thus, in this manner, the transfer function of the system can be synthesized (8, 24).

Care must be exercised in determining whether any zeros of the transfer function are in the right-half s plane. A system that has no open-loop zeros in the right-half s plane is defined as a minimum-phase system (8). A system that has open-loop zeros in the right-half s plane is a non-minimum-phase system. The stability is determined by the location of the poles and does not affect the designation of minimum or non-minimum phase.

The angular variation for poles or zeros in the right-half s plane is different from those in the left-half plane. For this situation, one or more terms in the transfer function have the form $1 - Ts$ and/or $1 \pm As \pm Bs^2$. As an example, consider the functions $1 + j\omega T$ and $1 - j\omega T$. The log magnitude plots of these functions are identical, but the angle plot for the former, as ω varies from 0 to ∞, goes from 0° to 90°, whereas for the latter it goes from 0° to −90°. Therefore, care must be exercised in interpreting the angle plot to determine whether any factors of the transfer function lie in the right-half plane.

CLOSED-LOOP FREQUENCY RESPONSE

Specifications for closed-loop system performance are often stated either in terms of the time response characteristics with a specified input or in terms of the frequency response. The time response figures of merit with a unit step input may include: peak overshoot, M_p; peak time, t_p; final value, y_{ss}; and settling time, t_s. Additional figures of merit that may be specified include: system type, m; gain constant K_m; and duplicating time, t_D (time to first reach the final value). For systems of higher than second order, an effective damping ratio ζ_{eff} may be used to compare the response to that of a simple second-order system. The settling time t_s is affected principally by the dominant closed-loop poles, with modifying effects due to the other poles and zeros. There is, therefore, an effective damping ratio ζ_{eff}.

An alternate or supplementary way of identifying performance specifications is in terms of frequency response characteristics. For a simple second-order system, frequency response characteristics consist of a set of figures of merit which are the maximum value M_m and the frequency ω_m at which the maximum value occurs. The relationships for a simple second-order system are

$$M_m = \frac{1}{2\zeta\sqrt{1-\zeta^2}}, \qquad \omega_m = \omega_n\sqrt{1-2\zeta^2} \tag{105}$$

These values are correlated to the peak value M_p and time response frequency of oscillation ω_d in the time domain by

$$M_p = 1 + \exp\left(-\frac{\zeta\pi}{\sqrt{1-\zeta^2}}\right), \qquad \omega_d = \omega_n\sqrt{1-\zeta^2} \tag{106}$$

Therefore, designing for given values of M_m and ω_m results in effective values of ζ_{eff} and ω_n. These, in turn, yield corresponding values of the time domain figures of merit which include M_p, ω_d, and, for a 2% settling time,

$$t_s = \frac{4}{\zeta\omega_n}, \qquad t_p = \frac{\pi}{\omega_n\sqrt{1-\zeta^2}} \tag{107}$$

When $\mathbf{G}(j\omega)$ is plotted in a polar (Nyquist) plot, there exists a family of constant M closed-loop contours, where $M = |C(j\omega)/R(j\omega)|$. These contours are a set of circles with specified centers and radii (8):

$$x_0 = -\frac{M^2}{M^2-1}, \qquad y_0 = 0, \qquad r_0 = \left|\frac{M}{M^2-1}\right| \tag{108}$$

Therefore, the gain constant K_m can be selected which makes the plot of $G(j\omega)$ tangent to a desired value M_m. In accordance with Eq. (105), this results in a corresponding set of values of ζ and ω_n. This produces a set of affiliated values of M_p, ω_d, t_s, and t_p, as given by Eqs. (106) and (107). The polar Nyquist plot therefore not only provides information on closed-loop stability, but also provides a design process for determining degree of stability. This degree of stability provides information on the time response characteristics.

The constant M circles on the polar plot can be transformed to the Nichols (Lm versus angle plot) plot. Then, adjusting the gain K_m raises or lowers the plot of $G(j\omega)$. This procedure can be used to make the plot of $\text{Lm}\, G(j\omega)$ tangent to the desired M curve on the Nichols chart. The frequency at the point of tangency determines the value of ω_m. Equations (105)–(107) determine the time response characteristics.

NONLINEAR STABILITY AND CONTROL

In this article the stability of dynamical systems is treated in a control theoretic context. Thus, the stability of feedback control systems is discussed. This is in the tradition of the classical Nyquist stability method in linear control theory covered in the preceding sections, where, based on information about the open-loop plant, the stability of the closed-loop feedback control system is being ascertained. In the remaining sections of this article the more realistic problem of the stability of constrained feedback control systems is considered. Specifically, the stability of closed-loop feedback control systems which consist of linear plants driven by actuation elements which are subject to hard saturation constraints is addressed. Obviously, under conditions of "small perturbations," the hard actuator constraints are not active and stability is then determined according to the "first method of Lyapunov" using conventional linear analysis methods. Attention is now given to the treatment of high amplitude maneuvers and large perturbations away from trim, namely, the treatment of the "stability in the large" of control systems comprised of linear plants and actuators which are subject to saturation.

Even though the primary interest is in stability, a complete analysis and a mathematically rigorous treatment of the stability problem in a control context is undertaken by embedding the stability problem into the more general framework of tracking control. Thus, the problem of tracking a zero exogeneous command signal naturally reduces to the investi-

gation of the feedback control system's stability. Moreover, actuator saturation is addressed by utilizing a nonlinear dual-loop control architecture, where (internally modified) reference signals are generated by the feedback controller, even in the absence of a nonzero exogeneous command signal. Indeed, the feedback controller responds to the exogeneous reference signal and to the current state measurement; and the ill effects of saturation caused by large excursions in the state can be mitigated by injecting into an inner-loop linear controller a modified (nonzero) reference signal that tries to pull the control system out of saturation. This situation, where the state excursions are large, can arise even in the case where the exogeneous reference signal vanishes for a period of time—in which case the feedback control system is, strictly speaking, a regulator. Finally, in the case where the feedback control system is nonlinear, the question of input/output [viz., bounded input, bounded output (BIBO)] stability becomes dominant. Whereas it is well known that asymptotically stable linear control systems are BIBO stable—that is, stability with respect to perturbations in the initial state also guarantees the input/output stability of the control system, and vice versa—the same cannot be said about nonlinear control systems. Hence, the response of the nonlinear control system to exogeneous reference signals needs to be considered and the broader issue of BIBO stability of tracking control systems must be addressed.

In the remaining sections of this article, the design of nonlinear tracking controllers for the mitigation of actuator saturation effects is discussed. Evidently, the development of a nonlinear controller synthesis methodology which addresses the conflicting requirements of high-amplitude dynamic reference signal tracking and regulation in the face of hard actuator displacement and rate constraints, actuator dynamics, and unstable open-loop plants is sorely needed. At the same time, in many applications (e.g., in flight control), full-state feedback can be assumed. Current analysis and design methods can address one, or some, of the above requirements, but, taken together, these requirements represent, at best, a difficult task, in particular with respect to satisfying the inherent tradeoffs between tracking control/regulation performance and closed-loop nonlinear system input/output and asymptotic stability. The objective here is to critically examine the state of the art and point the reader to a workable tracking control paradigm which, from the outset, also acknowledges actuator constraints.

The adverse effects of controller-induced state and/or control constraint violation on closed-loop system performance and stability are well documented for both stable and unstable open-loop plants, namely, degraded system performance, limit cycling, and unstable system response. In addition, in feedback control systems with open-loop unstable plants (e.g., in modern flight control) and where feedback is used for stabilization, saturation can cause the feedback loop to be opened and immediately induce instability—as opposed to more benign open-loop stable plants, where feedback does not play such a critical role. Moreover, in flight control, and for properly designed aircraft, actuator rate saturation usually occurs before actuator deflection saturation.

There is much interest with regard to constrained control in the recent literature, and actuator saturation is a topic of active research in control theory. Thus, numerous controller design and modification methodologies exist which address linear time-invariant (LTI) systems with state and/or control constraints (11,24–48). Much attention has been given to regulation and set-point control, as opposed to tracking control. Also, most of the work fails to address actuator rate saturation or, for that matter, actuator dynamics. Notable exceptions concerning tracking control are Refs. 30, 33, 34, 46, and 51, and rate saturation is explicitly considered in Refs. 30, 38, and 46. In most cases, open-loop stable plants are assumed and conditions are devised for global stability. In these investigations the compensator is allowed to send infeasible control signals to the actuator. In the landmark treatise (50), feedback control systems with amplitude constrained actuators and stable open-loop plants are considered. Necessary and sufficient frequency-domain conditions, which are a generalization of the Nyquist criterion, are given for global stability in the face of windup. Many of the proposed methodologies can be classified as anti-windup methods (26,35,44): The goal of anti-windup controllers is to avoid or delay actuator saturation-induced windup of linear dynamic compensation. While these methods may prevent or reduce windup and improve closed-loop system performance, they do not fully prevent constraint violation, namely, actuator saturation, which is the root cause of windup. Ad hoc anti-windup compensators (26,33,35,37,39,44,52), don't perform well in the case of open-loop unstable plants with actuator displacement and rate constraints. Finally, unstable open-loop plants are specifically addressed in Refs. 28, 30, 34, 46, and 51.

State and control constraint mitigation strategies based on saturation avoidance focus on constraint violation prevention. Recent work on saturation avoidance methods has produced several important results based on the concepts of output admissibility and static admissibility which are applicable to the problem of tracking control (11,29,30,34–49). Gilbert's discrete-time reference governor (DTRG) (31–32) is representative of these methods. Another interesting method is the linear quadratic tracking (LQT) concept (40–42), which uses a receding horizon optimal control paradigm to avoid downstream state and control constraint violations.

In this article, and motivated by problems in modern flight control, the synergetic consideration of stability and tracking control of open-loop unstable plants, with actuators subject to rate and amplitude saturation, using state feedback, is undertaken. Since Refs. 31, 32, and 40–42 are representative of the state of the art with respect to constrained tracking control, their results are reviewed here. A viable feedback control concept is presented in the section entitled "Control Concepts." The section entitled "Linear Quadratic Tracking (LQT)" discusses the LQT methodology for saturation avoidance and improved tracking performance developed in Refs. 40–42. The section entitled "Maximal Output Admissible Sets" discusses the concept of maximal output admissible sets which play a crucial role in the synthesis of BIBO tracking control systems. The section entitled "Discrete-Time Reference Governor" is devoted to Gilbert's DTRG, which is illustrated using a scalar example. In the section entitled "Static Admissibility and Invariance" the same scalar example is used to investigate more general invariance-based control methods with guaranteed BIBO stability and also yield improved tracking performance; some fine points are illustrated. A suboptimal approach is presented in the section entitled "Suboptimal Approach," which is followed by concluding remarks in the section entitled "Conclusion."

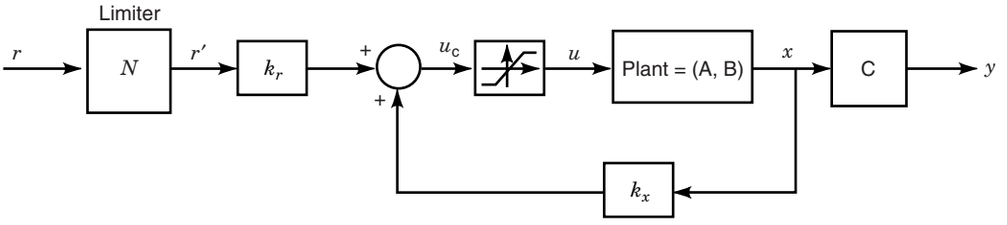

Figure 18. Current control system with open-loop limiter.

CONTROL CONCEPT

Figures 18–21 show examples of nonlinear control system architectures for linear systems with control constraints.

Figure 18 shows the typical nonlinear control system architecture currently used in (flight control) practice: A simple open-loop limiter is used to restrict the reference signal and help mitigate the adverse effects of actuator saturation-caused windup. Figure 19 is a general nonlinear feedback control system architecture where the nonlinear controller, G, computes the commanded control signal based on the current plant state and the current, and possibly past, values of the reference signal. The controller is cognizant of the saturation level of the downstream actuator which is located at the plant input. Hence, the controller-generated signal, u_c, may attain its limits, and by doing so the plant is driven to its full capacity. Figure 20 shows a specific example of a general nonlinear controller's architecture. A nonlinear dual-loop control system is envisaged where the controller's nonlinearity is confined to a feedback limiter, N, which performs a nonlinear scaling of the reference signal. The inner loop controller is linear and the nonlinear element N is contained in the outer loop, as illustrated in Fig. 20. Furthermore, the inner-loop linear controller is the result of a linear design, wherein it is assumed that the actuator operates in the linear regime. Hence, the inner loop is a linear control system. The nonlinear element N in the outer loop is a (pilot) command limiter which employs feedback. The nonlinear element N scales the reference signal such that the commanded control signal u_c does not violate the actuator imposed control constraints in the inner loop: The feedback (pilot) command limiter generates a modified reference signal r' which drives the inner loop in such a way that saturation in the downstream actuator is precluded, and strictly linear action ensues in the inner loop. The outer loop's nonlinearity N renders transparent the actuator's nonlinearity in the inner loop, and therefore the feedback control system shown in Fig. 20 is equivalent to the simpler feedback control system of Fig. 21, where the inner loop is linear. In conclusion, in Figs. 20 and 21 the linear inner loop controller is designed to meet small signal tracking and stability specifications, and the nonlinear element N in the outer loop is designed to yield good tracking performance and BIBO stability. Obviously, the design challenge is to maintain tracking performance and the BIBO stability guarantee of the dual loop nonlinear control system, for a relatively large class of exogeneous reference signals, and for a sufficiently large set of initial states.

Clearly, the control schemes shown in Figs. 20 and 21 are more advanced than the simple control scheme shown in Fig. 18, and are discussed in this article.

LINEAR QUADRATIC TRACKING (LQT)

It is convenient to address the actuator constrained tracking control problem using a time-domain, receding-horizon, optimal control formulation. Now, application of optimal control methods to the tracking problem requires a priori knowledge of the dynamic reference signal for all time. This also applies to linear quadratic (LQ) optimal control. Also, knowledge of the reference signal ahead in time is required, irrespective of whether the control signal is, or is not, constrained. Of course, this a priori knowledge does not exist, and in Refs. 40–42, tracking control is addressed using reference signal extrapolation and implementing a receding-horizon optimal control methodology in which the exogeneous (pilot) reference signal is predicted over the optimization horizon, based on current and near past pilot reference signal values. Specifically, polynomial extrapolation and interpolation is used to obtain the reference signal over the predetermined optimization horizon. Then, within each receding horizon window (RHW), the reference signal vector, $\hat{\mathbf{r}}$, is known, so that the LQ optimal control methodology may be applied to each of the finite-horizon control problems. The resulting finite-horizon LQ optimal control problem is solved within each RHW. In this way the indefinite-horizon control problem is broken up into an open-ended sequence of finite-horizon control problems. In a discrete-time formulation the optimization window is of length N (= the number of samples within each RHW).

Specifically, an inner-loop linear quadratic controller, designed for good small-signal performance, minimizes a quadratic cost functional over the optimization horizon N. Thus, suppose that the bare plant is

$$x(k+1) = Ax(k) + Bu(k), \quad x(0) = x_0, \qquad k = 0, 1, \ldots, N-1$$

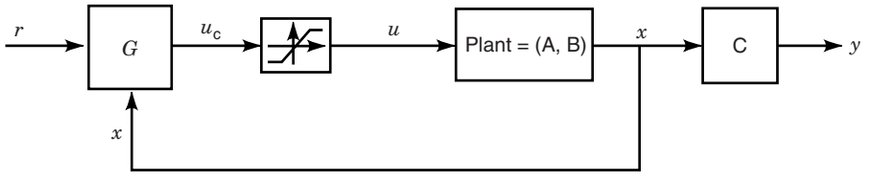

Figure 19. Nonlinear control system.

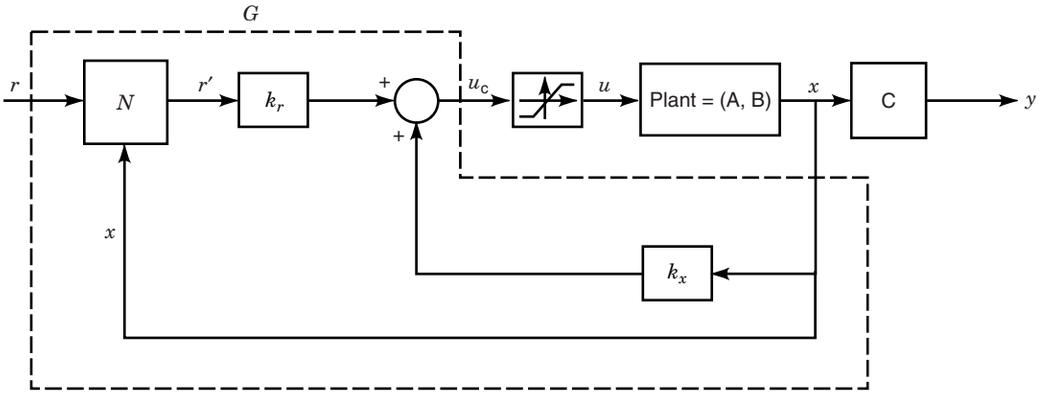

Figure 20. Two-loop control concept.

and the tracked output signal is

$$y(k+1) = Cx(k+1)$$

The cost functional is

$$J(\mathbf{u}) = \sum_{k=0}^{N} [Q(r(k+1) - y(k+1))^2 + Ru^2(k)]$$

The Q and R weights establish the tracking performance/control effort tradeoff in the inner (linear) loop and determine the "small signal" performance of the control system. Full-state feedback is assumed.

For a given reference sequence $\hat{\mathbf{r}}$, LQ optimal control returns the optimal control time history $\mathbf{u}^* = [u_0^*, u_1^*, \ldots, u_{N-1}^*]$, obtained within each RHW as a linear function of the complete reference vector $\hat{\mathbf{r}} = [r_1, \hat{r}_2, \ldots, \hat{r}_N]$, and the initial plant state, $\mathbf{x}_0$, where r_1 is the currently commanded reference signal. Now, polynomial extrapolation and interpolation yields $\hat{r}_2, \ldots, \hat{r}_N$, the reference signal's prediction, linear in the data. Specifically, if, for example, a simple ZOH extrapolation strategy is used, it is shown in Ref. 40 that the predicted reference vector $[\hat{r}_2, \ldots, \hat{r}_N]$ is linear in r_1; that is, $\hat{\mathbf{r}}$ is linear in r_1, which ultimately yields the optimal control sequence $[u_0^*, u_1^*, \ldots, u_{N-1}^*]$ linear in r_1, the current pilot-demanded reference signal, and x_0. Note that r_1 is the actual pilot demanded reference signal at time now, and is not a predicted value. Hence, in the current optimization window, LQ optimal control returns the optimal control time history

$$u^*(i) = k_{i_x} x_0 + k_{i_r} r_1, \qquad i = 0, 1, \ldots, N-1 \tag{109}$$

In particular, at time 0, at the start of the window, $u^*(0) = k_{0_x} x_0 + k_{0_r} r_1$. In the sequel, the notation used is

$$k_x \equiv k_{0_x}, \qquad k_r \equiv k_{0_r}$$

so that

$$u^*(0) = k_x x_0 + k_r r_1 \tag{110}$$

Both k_x and k_r are provided by the solution of the finite horizon LQ optimal control problem. The row vector $k_x^T \in R^n$ corresponds to the gain derived from the solution, over the finite optimization horizon, of the Riccati difference equation; this same Riccati equation is also associated with the solution of the finite horizon regulation problem. Moreover, one can deviate from the optimal tracking control solution in the RHW and instead, once k_x has been determined, choose the gain k_r such that asymptotic tracking of a step reference command r is enforced. Of course, if integral action is used to obtain a type-one system, then *asymptotic* tracking is achieved for any k_r.

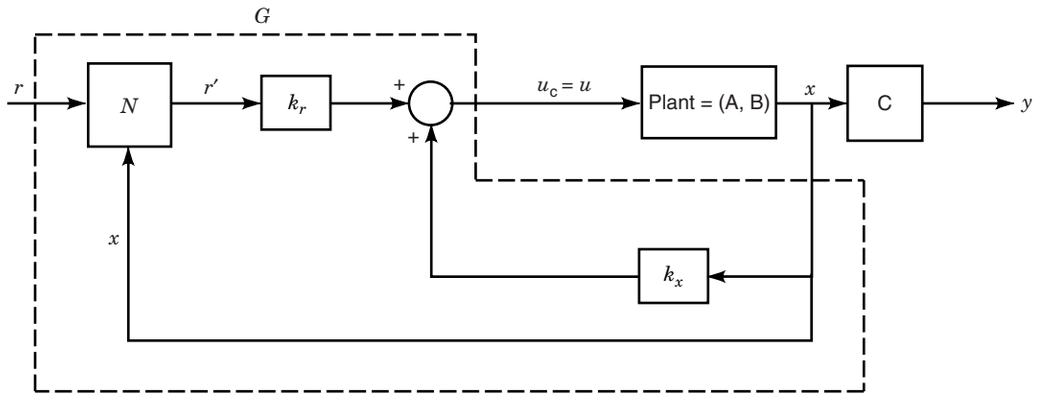

Figure 21. Saturation mitigation concept.

Furthermore, since the optimal control time history $\mathbf{u}^*$ is a linear function of r_1 and x_0, the actuator amplitude constraints of, say, $\pm u_{\max}$, can be readily transformed into constraints on the reference signal at time now, r_1: In the RHW

$$-u_{\max} \le k_{i_x} x_0 + k_{i_r} r_1 \le u_{\max}, \qquad i = 0, 1, \ldots, N-1$$

there is nothing that can be done about the plant's initial state x_0. Hence, for example, for $k_r > 0$ the above inequalities yield the explicit constraints on the exogeneous reference signal:

$$\max_{0\le i\le N-1}\left(\frac{-u_{\max} - k_{i_x} x_0}{k_{i_r}}\right) \le r_1 \le \min_{0\le i\le N-1}\left(\frac{u_{\max} - k_{i_x} x_0}{k_{i_r}}\right)$$

Next, differencing Eq. (109) yields an explicit expression for the actuator rate, namely,

$$\begin{aligned}\dot{u}_i &\equiv \frac{1}{\Delta T}(u_i - u_{i-1}) \\ &= \frac{1}{\Delta T}(k_{i_x} x_0 + k_{i_r} r_1 - k_{i-1_x} x_0 - k_{i-1_r} r_1) \\ &= \frac{1}{\Delta T}[(k_{i_x} - k_{i-1_x}) x_0 + (k_{i_r} - k_{i-1_r}) r_1]\end{aligned}$$

Similar to the derivation for actuator amplitude constraints, the above equations are manipulated to yield an additional explicit bound on the reference signal at time now, r_1,

$$\begin{aligned} r_1 &\ge \left\{ \max_{1\le i\le N-1}\left(\frac{-\Delta T \dot{u}_{\max} - (k_{i_x} - k_{i-1_x}) x_0}{k_{i_r} - k_{i-1_r}}\right), \frac{u^*_{-1} - k_x x_0 - \Delta T \dot{u}_{\max}}{k_r} \right\} \\ r_1 &\le \left\{ \min_{1\le i\le N-1}\left(\frac{\Delta T \dot{u}_{\max} - (k_{i_x} - k_{i-1_x}) x_0}{k_{i_r} - k_{i-1_r}}\right), \frac{u^*_{-1} - k_x x_0 + \Delta T \dot{u}_{\max}}{k_r} \right\} \end{aligned} \tag{111}$$

where u^*_1 is the control signal from the last window. Hence, actuator constraints are easily transformed into constraints on the current reference signal. Saturation avoidance is guaranteed, provided that the reference signal at time now, r_1, satisfies the above inequalities.

Actuator displacement and rate constraints impose constraints on the inner closed-loop's exogenous reference signal. Thus, an additional outer control loop is employed which performs a nonlinear modification of the reference signal from the pilot, such that the downstream actuator constraints in the inner loop are satisfied. Now, if r_1 causes the violation of any of the constraints, it may be optimally scaled such that the modified reference signal, r'_1, satisfies all constraints; here, "optimally scaled" means that in the outer loop r'_1 is chosen so that, for example, $|r_1 - r'_1|$ is minimized, subject to r'_1 satisfying the actuator constraints.

In accordance with the receding-horizon modus operandi, at each time step the optimization window is shifted forward in time, a new state is attained, and upon receipt of a new reference signal from the pilot the reference signal extrapolation algorithm is reapplied and the open-loop LQ optimal control problem is solved from the beginning. Hence, r_1 (or r'_1) is the only reference value actually tracked; and within each RHW, u^*_0 is the only control which is actually applied to the system, so that feedback action is achieved. Now, the window's "local" time instant 0 corresponds to the current time k. Thus, at time k, the optimal control signal u^*_k satisfies

$$u^*(k) = k_x x(k) + k_r r(k+1)$$

and since the nonlinear element N in the outer loop causes the exogeneous reference signal r_{k+1} to be replaced by r'_{k+1}, which is then sent to the inner loop, the actual optimal control signal is

$$u^*(k) = k_x x(k) + k_r r'(k+1) \tag{112}$$

Many of the constraints on r_{k+1} are induced by constraints on subsequent elements of the predicted reference vector ($\hat{r}_{k+2}$, $\hat{r}_{k+3}$, . . ., $\hat{r}_{k+N}$). But these reference signals are not actually realized, so enforcement of these "downstream" constraints may lead to an overly conservative scaling of r_{k+1}. Thus, the designer may wish to specify feasibility criteria which do not necessarily include all the constraints in the RHW. An important special case entails the enforcement of actuator constraints at time now only, whereupon the inequalities from above are reduced to

$$\begin{aligned} r_1 &\ge \max\left(\frac{-u_{\max} - k_{i_x} x_0}{k_{i_r}}, \frac{u^*_{-1} - k_x x_0 - \Delta T \dot{u}_{\max}}{k_r}\right) \\ r_1 &\le \min\left(\frac{-u_{\max} - k_{i_x} x_0}{k_{i_r}}, \frac{u^*_{-1} - k_x x_0 - \Delta T \dot{u}_{\max}}{k_r}\right) \end{aligned} \tag{113}$$

Finally, if r_{k+1} satisfies the applicable constraints, then N is transparent and clearly $r'_{k+1} = r_{k+1}$, that is, tracking performance during "small-signal" operation is not sacrificed. Then, the closed-loop linear system, which translates the exogeneous r command into aircraft responses, is given by

$$\begin{aligned} x(k+1) &= A_{cl} x(k) + B_{cl} r(k+1) \\ y(k+1) &= C x(k+1) \end{aligned}$$

where $A_{cl} = A + Bk_x$, $B_{cl} = Bk_r$, and the actuator displacement is

$$u(k) = k_x x(k) + k_r r(k+1)$$

In general, the nonlinear LQT control law depends on (1) the choice of the Q and R weights for small-signal performance, (2) the reference vector prediction method, (3) the selection of the applicable feasibility criteria for saturation avoidance, and (4) the modified reference signal optimization criterion.

So far, only saturation avoidance was considered for modified reference signal feasibility determination. BIBO stability is only guaranteed under certain additional restrictive conditions. Thus, the design for guaranteed BIBO stability of the control system shown in Figs. 20 and 21 is now undertaken.

MAXIMAL OUTPUT ADMISSIBLE SETS

In Ref. 32 the concept of maximal output admissible set is employed to develop reference governors N for discrete-time

systems. The discrete-time reference governor (DTRG) developed in Ref. 32 is a nonlinear dynamical element, that is, it is a first-order lag filter with a variable bandwidth parameter,$\lambda \in [0, 1]$, which scales the reference signal's increments so that the controlled system's constraints are not violated. These constraints characterize the maximal output admissible set. By forming a modified reference signal such that the discrete-time system's state update satisfies these constraints, both saturation avoidance and BIBO stability are enforced: BIBO stability also requires that the maximal output admissible set be bounded. In the sequel, the DTRG-based approach to tracking control is discussed. More importantly, the construction of the maximal output admissible set is presented. This is a critical step in the development of more general BIBO tracking control systems, in both the discrete-time and the continuous-time settings.

The LTI discrete-time system

$$\begin{gathered} x(t+1) = Ax(t) + Bu(t), \\ x(0) = x \in \Re^n, \\ u(t) \in \Re^m, \quad t = 0, 1, 2, \ldots (= I^+) \end{gathered} \tag{114}$$

is considered. Initially, the regulation problem is exclusively analyzed and the prespecified small-signal linear-state feedback control law is

$$u(t) = k_x x(t)$$

Also, assume there are constraints on both the state and control vectors, including linear combinations of them. Then, with appropriate choices of matrices C and D, along with a set Y, these rather general constraints may be represented as

$$y(t) = Cx(t) + Du(t) \in Y \subset \Re^p \tag{115}$$

Let $A_{cl} = A + Bk_x$ and $C_{cl} = C + Dk_x$. Then, Eqs. (114) and (115) become

$$\begin{gathered} x(t+1) = A_{cl}x(t), \quad x(0) = x, \qquad t \in I^+ \\ y(t) = C_{cl}x(t) \in Y \subset \Re^p \end{gathered} \tag{116}$$

Hence, the saturation avoidance problem has been transformed into a feasibility problem consisting of an unforced LTI discrete-time system with an output constraint. In Ref. 31 the maximal output admissible set associated with system (116) is defined as the set of *all* initial states $x \in \Re^n$ such that the unforced closed-loop linear system's response does not violate the system output constraints for all $t \in I^+$. Thus, the maximal output admissible set is the largest set of initial states, $x \in \Re^n$, that is positively invariant w.r.t. system [see Eq. (116)]. Evidently, the maximal output admissible set is

$$O_\infty(A_{cl}, C_{cl}, Y) = \{x \in \Re^n : C_{cl}A_{cl}^t x \in Y \quad \forall\, t \in I^+\} \tag{117}$$

Example 1. The Output Constraint Set Y = {0}. The maximal output admissible set is then the subspace of unobservable states that corresponds to the pair (A_{cl}, C_{cl}), namely,

$$O_\infty = (\mathbf{N}(C_{cl}))_{A_{cl}}$$

This is the largest subspace contained in the null space of the matrix C_{cl} which is invariant under A_{cl}.

Example 2. The Set Y is a Cone. This is the situation where there are one-sided control constraints, for example, $0 < u(t)$—in which case C = nonnegative orthant in $\Re^m \equiv$ set of vectors in $\Re^m$ with non-negative entries. Then

$$O_\infty(A_{cl}, C_{cl}, Y) = \{0\}$$

iff

1. The pair (A_{cl}, S) is observable, where the n-column matrix S is such that $Sx = 0$ implies that $C_{cl}x \in Y$, and
2. A_{cl} does not have an eigenvector v which corresponds to a nonnegative eigenvalue of A_{cl} such that $C_{cl}v \in Y$.

For a proof, see, for example, Ref. 45. Of major interest is the case where the constraint set Y is a polyhedron, namely,

$$Y = \{y \in \Re^p : f_i(y) \le 0, \quad i = 1, 2, \ldots, s\} \tag{118}$$

That is, the s functions $f_i(y)$, $f_i : \Re^p \to \Re$ are linear in y, with $f_i(0) \le 0$. Obviously

$$\begin{aligned} O_\infty(A_{cl}, C_{cl}, Y) = \{x \in \Re^n : f_i(C_{cl}A_{cl}^t x) \le 0, \\ i = 1, 2, \ldots, s, \text{ and } t \in I^+\} \end{aligned} \tag{119}$$

Although Y is a polyhedron, Eqs. (117) and (119) each represent an infinite number of constraints. However, if there exists a finite $t_i^* \in I^+$ for each inequality in (119) for which the constraints associated with the ith inequality constraint are inactive for $t > t_i^*$, let $t^* = 1 \le i \le s \max t_i^*$. Then $O_\infty(A_{cl}, C_{cl}, Y)$ is characterized by a finite number ($\le st^*$) of inequality constraints—in this case, $O_\infty(A_{cl}, C_{cl}, Y)$ is said to be finitely determined—and determination of whether a particular initial state vector, x, is an element of $O_\infty(A_{cl}, C_{cl}, Y)$ involves evaluation of a finite set of linear inequalities. Moreover, it may transpire that the constraints associated with one or more of the inequalities, f_i, that define Y are inactive for all $t \in I^+$. Thus, let S^* denote the set of inequality constraints that are active for some $t \in I^+$, namely, $S^* \subset \{1, 2, \ldots, s\}$. Then, $O_\infty(A_{cl}, C_{cl}, Y)$ may be written as

$$\begin{aligned} O_\infty(A_{cl}, C_{cl}, Y) = \{x \in \Re^n : f_i(C_{cl}A_{cl}^t x) \le 0, \\ t = 0, 1, \ldots, t_i^*, \text{ and } i \in S^*\} \end{aligned} \tag{120}$$

In the discrete-time case under consideration, sufficient conditions for the existence of a finite t^*, namely, sufficient conditions for the finite determination of O_∞, are as follows: (1) A_{cl} is asymptotically stable, (2) $0 \in \text{int}(Y)$, (3) Y is bounded, and (4) the pair (A_{cl}, C_{cl}) is observable. Here, $\text{int}(\cdot)$ denotes the interior of the set.

The concept of maximal output admissible sets is only applicable to unforced systems. Thus, it is not directly applicable to the tracking problem without some modification. In Ref. 32 the concept of maximal output admissible sets is adapted for use with the tracking problem by using a nonlinear element in the DTRG which has first-order dynamics; that is, a first-order lag filter with a variable bandwidth parameter $\lambda \in [0, 1]$ is used to prefilter the exogenous reference signal. The closed-loop system state vector is then augmented with the prefilter state, namely, the modified reference signal, r'. The end result is an augmented system whose exogenous input

can be turned off by setting $\lambda = 0$. Thus, consider the tracking control problem

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t), \qquad x(t) \in \Re^n, u(t) \in \Re^m \\ u(t) &= k_x x(t) + k_r r(t) \end{aligned} \tag{121}$$

with state and control constraints

$$y(t) = Cx(t) + Du(t) \in Y \subset \Re^p \tag{122}$$

and where the control constraint set Y is given by Eq. (118). The ensuing closed-loop system is

$$\begin{aligned} x(t+1) &= A_{cl}x(t) + B_{cl}r(t) \\ y(t) &= C_{cl}x(t) + D_{cl}r(t) \in Y \subset \Re^p \end{aligned} \tag{123}$$

where, as before, $A_{cl} = A + Bk_x$, $C_{cl} = C + Dk_x$, and $B_{cl} = Bk_r$, $D_{cl} = Dk_r$. To transform this problem into one which allows use of the concept of maximal output admissible sets, the exogenous reference signal, r, is prefiltered by the first-order lag filter given by

$$r'(t+1) = r'(t) + \lambda(r(t), x_g(t))\big(r(t) - r'(t)\big) \tag{124}$$

where $\lambda(r(t), x_g(t)) \in [0, 1]$; and in Eq. (119) $r(t)$ is replaced by the filter's output $r'(t)$ (the modified reference signal). Then, the augmented state is

$$x_g = \begin{bmatrix} r' \\ x \end{bmatrix} \tag{125}$$

and the augmented system dynamics and the output constraints are given by

$$\begin{aligned} x_g(t+1) &= A_g x_g(t) + B_g \lambda(r(t), x_g(t))\big(r(t) - [I \quad 0]x_g(t)\big) \\ y(t) &= C_g x_g(t) \in Y \subset \Re^p \end{aligned} \tag{126}$$

where

$$A_g = \begin{bmatrix} I & 0 \\ B_{cl} & A_{cl} \end{bmatrix}, \qquad B_g = \begin{bmatrix} I \\ 0 \end{bmatrix}, \qquad C_g = [D_{cl} \quad C_{cl}]$$

From Eq. (124) notice that if $\lambda(r(t), x_g(t)) = 1$, the exogenous reference signal, $r(t)$, is passed through unmodified, but with a one time-step delay. More importantly, if $\lambda(r(t), x_g(t)) = 0$ the current modified reference signal, $r'(t)$, remains unchanged, and Eq. (126) becomes an unforced system with output constraints. Thus, if at time $t = \tau$ and for all "initial states" $x_g(\tau) \in O_\infty(A_g, C_g, Y)$, $\lambda(r(\tau), x_g(\tau))$ can be chosen such that the updated state $x_g(\tau + 1) \in O_\infty(A_g, C_g, Y)$, it is possible to guarantee that the controlled system's state and control constraints will not be violated in the future. Moreover, assume that $O_\infty(A_g, C_g, Y)$ is bounded. Then because $\lambda = 0$ is always an option and because the new reference signal is chosen such that it does not remove the state of the plant from the compact set of admissible states, the BIBO stability of tracking control systems which employ the DTRG is guaranteed.

If Y is given by Eq. (118) and $O_\infty(A_g, C_g, Y)$ is finitely determined—that is, it is characterized by a finite set of linear inequality constraints—then at each time-step the upper limit imposed on the scalar $\lambda(r(t), x_g(t))$ by each linear inequality constraint is given by a simple algebraic formula. Thus, at each time-step $\lambda(r(t), x_g(t)) \in [0, 1]$ may be chosen so that it satisfies the minimum of the upper limits and is easily computed on-line. Hence, a tracking controller could easily be synthesize; moreover, the output vector is guaranteed to be bounded.

Unfortunately, application of the concept of maximal output admissible sets to the tracking control problem generally results in an augmented dynamics matrix A_g which is only Lyapunov stable. Thus, $O_\infty(A_g, C_g, Y)$ is generally not finitely determined. The reason $O_\infty(A_g, C_g, Y)$ is not finitely determined stems from the fact that the unforced response of a Lyapunov stable linear system does not decay to the origin, unlike that of an asymptotically stable linear system. In fact, the unforced response of the Lyapunov stable linear system does converge to an a priori unknown equilibrium point. In general, the equilibrium point will not be reached in a finite number of time-steps, and thus t^* cannot be bounded. The following will be useful in resolving this problem.

The set of statically admissible inputs, R_s, for the system defined by Eq. (123), is the set of constant inputs for which the associated equilibrium point is admissible. That is,

$$R_s = \{r \in \Re^m : y_{ss} \in Y\}$$

where y_{ss} is the steady-state response. $y_{ss} = H_o r$, where $H_o = C_{cl}(I - A_{cl})^{-1} B_{cl} + D_{cl}$. Hence, an equivalent expression for R_s is

$$R_s = \{r \in \Re^m : H_o r \in Y\} \tag{127}$$

The set of statically admissible states, X_s, for the closed-loop system defined by Eq. (123), is the set of initial states for which there exists a statically admissible reference signal such that the ensuing trajectory does not violate the system's state and control constraints for all time. That is,

$$X_s = \{x \in \Re^n : \exists\, r \in R_s \quad s.t. \quad y(t) \in Y \quad \forall\, t \in I^+\} \tag{128}$$

For the augmented system defined by Eq. (126), the set of statically admissible states, X_{gs}, is

$$X_{gs} = O_\infty(A_g, C_g, Y) \tag{129}$$

Similar to the maximal output admissible set, both X_s and X_{gs} are positively invariant sets. X_s is the projection of X_{gs} onto the plane $r' = 0$. Now, recall that the reason $O_\infty(A_g, C_g, Y)$ is not finitely determined is that an infinite number of time-steps is required to reach the unknown equilibrium point. However, if it is somehow known that the unforced system given by Eq. (126) will converge to an equilibrium point $x_g \in \text{int}(X_{gs})$, then only the transient response for a finite number of time-steps needs to be considered, until the peak-to-peak magnitude of constrained quantity oscillations decay sufficiently. If r' is restricted to values in $\text{int}(R_s)$ and if $x(0) \in X_s$, then $x(t) \in X_s$ and $x_g(t) \in \text{int}(X_{gs})$ for all $t \in I^+$. Moreover, the set $\text{int}(X_{gs}) = \text{int}(O_\infty(A_g, C_g, Y))$ is finitely determined.

Now, define $Y(\epsilon) \subset Y$ by $Y(\epsilon) = \{y : f_i(y) \le -\epsilon,\ i = 1, \cdots, s\}$, where $0 < \epsilon < \min\{-f_j(0): i = 1, \cdots, s\}$. Then, to restrict r' to values in $\text{int}(R_s)$, which results in $x_g \in \text{int}(X_{gs})$, append

the additional constraints

$$f_i([H_o \quad 0]x_g) \le -\epsilon, \qquad i = 1, \ldots, s \tag{130}$$

to those that define $O_\infty(A_g, C_g, Y)$. Denoting $\text{int}(X_{gs})$ by X_{gs}^ϵ yields

$$X_{gs}^\epsilon = \{x_g \in \Re^{n+m}: f_i(C_g A_g^t x_g) \le 0, t = 0, \ldots, t_i^*, i \in S^*, \text{ and } f_j([H_o \quad 0]x_g) \le -\epsilon, j = 1, \ldots, s\} \tag{131}$$

Now, X_{gs}^ϵ is characterized by a finite set of inequality constraints. Moreover, the upper limit imposed on $\lambda(r(t), x_g(t)) \in [0, 1]$ by each inequality, so that $x_g(t + 1) \in X_{gs}^\epsilon$, is given by a simple formula. Thus, if $x_g(0) \in X_{gs}^\epsilon$, and at each time-step we choose $\lambda(r(t), x_g(t)) \in [0, 1]$ such that it satisfies the minimum of the upper limits imposed by all inequality constraints in Eq. (131), then $x_g(t) \in X_{gs}^\epsilon$ for all $t \in I^+$.

The finite determination of the maximal statically admissible (invariant) set is exclusively an artifact of discrete-time dynamics. The maximal output admissible invariant sets of continuous-time systems are *not* polyhedral.

DISCRETE-TIME REFERENCE GOVERNOR

The concept of a statically admissible set and the discrete-time reference governor (DTRG) are illustrated. Specifically, a tracking controller is synthesized for a constrained scalar control system. The continuous-time control system is given by

$$\dot{x} = ax + bu, \quad x(0) = x_o, \quad -1 \le u \le 1 \tag{132}$$

and the small-signal, linear control law is

$$u = k_x x + k_r r' \tag{133}$$

Combining Eqs. (132) and (133) gives

$$\dot{x} = a_{cl} x + b k_r r', \qquad a_{cl} = a + b k_x \tag{134}$$

Now, choose k_r such that asymptotic tracking is enforced, namely,

$$k_r = -\frac{a}{b} - k_x = -\frac{a_{cl}}{b}$$

Then, the inner loop in Figs. 20 and 21 is

$$\dot{x} = a_{cl}(x - r') \tag{135}$$

and

$$u = k_x x - \frac{a_{cl}}{b} r' \tag{136}$$

Application of the DTRG to the system represented by Eqs. (132) and (133) requires the consideration of the equivalent discrete-time closed-loop system. For example, with $a = 2$, $b = 1$, $k_x = -3$, $k_r = 1$, and a sampling interval of $T = 0.001$ s, the equivalent discrete-time closed-loop system and the applicable constraints are

$$x(k + 1) = a_d x(k) + b_d r'(k) \tag{137}$$

$$y(k) = u(k) = cx(k) + dr'(k) \in Y \subset \mathscr{R} \tag{138}$$

$$Y = \{y \in \Re : f_i(y) \le 0, i = 1, 2\} \tag{139}$$

where $a_d = 0.999$, $b_d = 9.995 \times 10^{-4}$, $c = k_x = -3$, $d = k_r = 1$, $f_1(y) = y - 1$, and $f_2(y) = -y - 1$. $r'(k)$ is the feasible reference signal.

The reference governor is

$$r'(k + 1) = r'(k) + \lambda(k)[r(k) - r'(k)] \tag{140}$$

Combining the closed-loop system and reference governor dynamics results in the augmented second-order system

$$x_g(k + 1) = A_g x_g(k) + B_g \lambda(k)(r(k) - [1 \quad 0]x_g(k)) \tag{141}$$

$$y(k) = C_g x_g(k) \in Y \tag{142}$$

where the augmented state is

$$x_g(k) = \begin{bmatrix} r'(k) \\ x(k) \end{bmatrix}$$

and

$$A_g = \begin{bmatrix} 1 & 0 \\ 9.995 \times 10^{-4} & 0.999 \end{bmatrix}$$

$$B_g = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad C_g = [1 \quad -3]$$

Now the maximal output admissible set for the augmented system is concerned with the homogeneous system [$\lambda(k) = 0$] and is defined as

$$O_\infty = \{x \in \Re : f_i(C_g A_g^k x_g) \le 0, k = 0, \ldots, k_i^*, i \in S^*\} \tag{143}$$

where the constraints, $f_i(C_g A_g^k x_g) \le 0$, are inactive for $k > k_i^*$ and $i \notin S^* \subset \{1, 2\}$. Since the augmented system is only Lyapunov stable, k_i^* may be unbounded, so a finitely determined approximation, O_∞^ϵ, to the maximal output admissible set is needed. As noted in the section entitled "Maximal Output Admissible Sets," O_∞^ϵ is obtained by restricting the feasible input, r', to values that are statically admissible with respect to the reduced constraint set $Y(\epsilon) = \{y : f_i(y) \le -\epsilon, i = 1, 2\}$, where $0 < \epsilon < \min\{-f_i(0): i = 1, 2\}$. Thus, the additional constraint is added:

$$r' \in W_0^\epsilon = \{r' : H_0 r' \in Y(\epsilon)\} \tag{144}$$

where

$$H_0 = d + c(1 - a_d)^{-1} b_d = -1.9985 \tag{145}$$

Then,

$$O_\infty^\epsilon = \{x \in \Re : f_i(C_g A_g^k x_g) \le 0, k = 0, \ldots, k_i^*, i \in S^*, f_j([H_0 \quad 0]x_g) \le -\epsilon, j = 1, 2\} \tag{146}$$

and k_i^* is guaranteed to be bounded (32). Notice that O_∞^ϵ is composed of two sets of constraints. The first set ($f_i(C_g A_g^k x_g) \le$

$0, k = 0, \ldots, k_i^*, i \in S^*$) deals with saturation of the transient response, and the second set ($f_j([H_0, 0]x_g) \le -\epsilon, j = 1, 2$) deals with saturation in steady state. Now, Algorithm 3.2 of Ref. 31 may be used to determine k_i^* and S^*. In this case S^* is empty. That is, the inequalities associated with the transient saturations are inactive. This should be expected because the closed-loop system of Eqs. (137) and (138) is a stable first-order (over damped) system. Thus, Eq. (146) becomes

$$O_\infty^\epsilon = \{f_j([H_0 \quad 0]x_g) \le -\epsilon, j = 1, 2\} \tag{147}$$

Now, with $\epsilon = 0.05$, Eqs. (3.10) and (3.15) of Ref. 32 may be used to generate $\lambda(k)$, namely,

$$\lambda(k) = \min\{\alpha_1(k), \alpha_2(k)\} \tag{148}$$

where

$$\alpha_1(k) = \begin{cases} \min\left\{1, \dfrac{0.95 - H_0 r'(k)}{H_0(r(k) - r'(k))}\right\} & \text{if } H_0(r(k) - r'(k)) > 0 \\ 1, & \text{if } H_0(r(k) - r'(k)) \le 0 \end{cases} \tag{149}$$

and

$$\alpha_2(k) = \begin{cases} \min\left\{1, \dfrac{0.95 + H_0 r'(k)}{H_0(r'(k) - r(k))}\right\} & \text{if } H_0(r'(k) - r(k)) > 0 \\ 1 & \text{if } H_0(r'(k) - r(k)) \le 0 \end{cases} \tag{150}$$

In this case, the end result of Eqs. (148)–(150) is to limit the feasible reference signal such that

$$-0.47536 \le r' \le 0.47536 \tag{151}$$

While the above reference governor avoids saturation and affords BIBO stability, it is somewhat conservative in that it restricts the feasible reference signal, r', such that at all times it is always statically admissible. The BIBO stable tracking controller developed in Refs. 40–42 employs a static nonlinearity N; it is nevertheless related to Gilbert's DTRG (32) in that it also includes "static admissibility" of the modified reference signal in the feasibility criteria. However, there are certainly cases where a modified exogenous reference input that is not statically admissible over a finite time interval would not necessarily result in saturation. Thus, as discussed in the next section, the achievable tracking performance can be enhanced, while at the same time the constraint violation is avoided and BIBO stability is guaranteed.

STATIC ADMISSIBILITY AND INVARIANCE

The scalar constrained control system is revisited and the application of statically admissible and more general (maximal output admissible) invariant sets to achieve saturation avoidance and BIBO stability enforcement is demonstrated. The construction of these invariant sets is a crucial step in the design of tracking control laws. Insights into the critical role played by the above-mentioned invariant sets permit focusing on design for tracking performance, while at the same time guaranteeing the BIBO stability of the control system. Moreover, the DTRG is now replaced by a static nonlinearity N whose synthesis is outlined. The flexibility afforded by this less conservative and more general approach makes it possible to consider two separate tracking control concepts. The first tracking control concept is $\min_{r'}|r - r'|$ subject to the control constraints. The second control concept for the selection of r' attempts to drive the output x to r as quickly as possible, subject to the control constraints. A continuous-time derivation is given.

Control Concept 1: Choose r' such that $|r - r'|$ is minimized subject to the control constraints. Assuming $k_x < 0$, $a_{cl} < 0$, and $b > 0$ results in the explicit (nonlinear) control law

$$r'(x, r) = \begin{cases} r & \text{if } \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} \le r \le \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} \\ \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} & \text{if } r \ge \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} \\ \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} & \text{if } r \le \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} \end{cases} \tag{152}$$

and substituting Eqs. (152) into (135) results in the closed-loop system

$$\dot{x} = \begin{cases} a_{cl}(x - r) & \text{if } \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} \le r \le \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} \\ ax + b & \text{if } r \ge \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} \\ ax - b & \text{if } r \le \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} \end{cases} \tag{153}$$

The saturation avoidance control law (152) does not guarantee BIBO stability in the case of an open-loop unstable plant. However, the desired BIBO stability can be obtained through an additional invariance requirement.

A bounded set $X_I \subset X = \{x: x \in \Re\}$ is invariant with respect to the system given by Eqs. (132), (133), and (153) if and only if on the boundary of X_I, $x\dot{x} < 0$. Thus, if a bounded invariant set is characterized as

$$X_I = \left\{x \in \Re : \exists\, r' \text{ such that } -1 \le u = k_x x - \frac{a_{cl}}{b}r' \le 1 \text{ and } x\dot{x} \le 0\right\}$$

then, by restricting x to $x \in X_I$, BIBO stability is guaranteed. First, consider the case of an open-loop unstable plant, $a > 0$. For $(bk_x/a_{cl})x + b/a_{cl} \le r \le (bk_x/a_{cl})x - b/a_{cl}$, the result is $r' = r$, and $\dot{x} = a_{cl}(x - r)$. Hence, $r < x$ results in $\dot{x} < 0$, and $r > x$ results in $\dot{x} > 0$. Second, for $r \ge (bk_x/a_{cl})x - b/a_{cl}$ we have $\dot{x} = ax + b$. In this case $x > -b/a$ results in $\dot{x} > 0$, and $x < -b/a$ results in $\dot{x} < 0$. Finally, for $r \le (bk_x/a_{cl})x + b/a_{cl}$ we have $\dot{x} = ax - b$. Therefore, $x > b/a$ results in $\dot{x} > 0$, and $x < b/a$ results in $\dot{x} < 0$. This is summarized in Fig. 22, where the directions of the arrows represent the sign of $\dot{x}$ in the Cartesian product space, V, defined by

$$V = \{v \in \Re^2 : v = [r, x]^T, r, x \in \Re\}$$

From Fig. 22 and the above discussion it is clear that if the control law allows $|x| > b/a$ the system will diverge due to the constraint on the control signal, u. Also, if the system ever

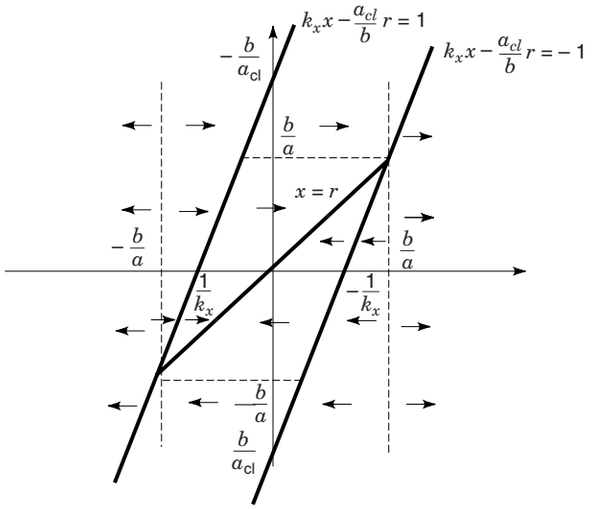

Figure 22. Cartesian product space, V, showing invariant region for the scalar system.

achieves either of the equilibrium points $v = [\pm b/a, \pm b/a]^T$, it becomes "stuck". Thus, in addition to control law (152), there must be a limit on x such that

$$x \in X_I = \left\{ x \in \Re : -\frac{b}{a} + \epsilon \le x \le \frac{b}{a} - \epsilon \right\} \tag{154}$$

then saturation will be avoided, and the system output will be bounded. Hence, the control law [see Eq. (152)] is modified as follows:

$$r'(x,r) = \begin{cases} r & \text{if } \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} \le r \le \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} \\ & \text{and } -\dfrac{b}{a} + \epsilon \le x \le \dfrac{b}{a} - \epsilon \\ \dfrac{b}{a} - \epsilon & \text{if } x > \dfrac{b}{a} - \epsilon \text{ and } r \ge \dfrac{b}{a} - \epsilon \\ -\dfrac{b}{a} + \epsilon & \text{if } x < -\dfrac{b}{a} + \epsilon \text{ and } r \le -\dfrac{b}{a} + \epsilon \\ \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} & \text{if } r > \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} \text{ and } x \le \dfrac{b}{a} - \epsilon \\ \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} & \text{if } r \le \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} \text{ and } x \ge -\dfrac{b}{a} + \epsilon \end{cases} \tag{155}$$

Now, under the explicit (nonlinear) control law [see Eq. (155)], $x \in X_I$ is enforced, the system will not become "stuck" at $x = \pm b/a$, and saturation avoidance plus invariance in a bounded subset of the state space yield the BIBO stability guarantee.

A similar analysis for the case of an open-loop stable plant ($a < 0$) shows that the saturation avoidance set is $X_I = \Re^1$. Hence, the simpler control law [see Eq. (152)] yields *global* BIBO stability in this case.

Control Concept 2: Choose r' so as to either maximize or minimize $\dot{x}$, based on the sign of $r - x$, and subject to the control constraints, therefore maximizing the instantaneous reduction in tracking error. Also, if $r = x$, then minimize $\dot{x}^2$, subject to the control constraints in Eq. (133). Thus,

$$\begin{aligned} &\text{If } r > x, \text{ then } r' \max(\dot{x}) \\ &\text{If } r < x, \text{ then } r' \min(\dot{x}) \\ &\text{If } r = x, \text{ then } r' \min(\dot{x}^2) \end{aligned}$$

From Eq. (135) and the assumption that $a_{cl} < 0$, this results in

$$\begin{aligned} &\text{If } r > x, \text{ then } \max r' \\ &\text{If } r < x, \text{ then } \min r' \\ &\text{If } r = x, \text{ then } \min\ (x - r')^2 \\ &\text{subject to } \frac{bk_x}{a_{cl}}x + \frac{b}{a_{cl}} \le r' \le \frac{bk_x}{a_{cl}}x - \frac{b}{a_{cl}} \end{aligned} \tag{156}$$

Hence, the explicit control law is finally obtained

$$r' = \begin{cases} \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}}, & \text{if } r > x \\ \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}}, & \text{if } r < x \\ r & \text{if } r = x \text{ and } -\dfrac{b}{a} \le x \le \dfrac{b}{a} \\ \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} & \text{if } r = x \text{ and } x < -\dfrac{b}{a} \\ \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} & \text{if } r = x \text{ and } x > \dfrac{b}{a} \end{cases} \tag{157}$$

and substituting Eq. (157) into Eq. (135) results in the "closed-loop" system,

$$\dot{x} = \begin{cases} ax + b & \text{if } r > x \\ ax - b & \text{if } r < x \\ 0 & \text{if } r = x, \text{ and } -\dfrac{b}{a} \le x \le \dfrac{b}{a} \\ ax + b & \text{if } r = x, \text{ and } x < -\dfrac{b}{a} \\ ax - b & \text{if } r = x, \text{ and } x > \dfrac{b}{a} \end{cases} \tag{158}$$

If the open-loop system is stable (i.e., $a < 0$), then control law (157) results in a globally BIBO stable closed-loop system. However, if the open-loop system is unstable ($a < 0$), then, as before, x must be restricted such that $x \in X_I$, where X_I is given by Eq. (154). In this case, control law [see Eq. (157)] is modified as follows:

$$r' = \begin{cases} \dfrac{bk_x}{a_{cl}}x - \dfrac{b}{a_{cl}} & \text{if } r > x, \text{ and } x \le \dfrac{b}{a} - \epsilon \\ \dfrac{bk_x}{a_{cl}}x + \dfrac{b}{a_{cl}} & \text{if } r < x, \text{ and } x \ge -\dfrac{b}{a} + \epsilon \\ r, \quad r = x & \text{if } -\dfrac{b}{a} + \epsilon \le x \le \dfrac{b}{a} - \epsilon \\ \dfrac{b}{a} - \epsilon & \text{if } r > x, \text{ and } x > \dfrac{b}{a} - \epsilon \\ -\dfrac{b}{a} + \epsilon & \text{if } r < x, \text{ and } x < -\dfrac{b}{a} + \epsilon \end{cases} \tag{159}$$

Now, $x \in X_I$ is enforced, the system will not become "stuck" at $x = \pm b/a$, and BIBO stability is achieved.

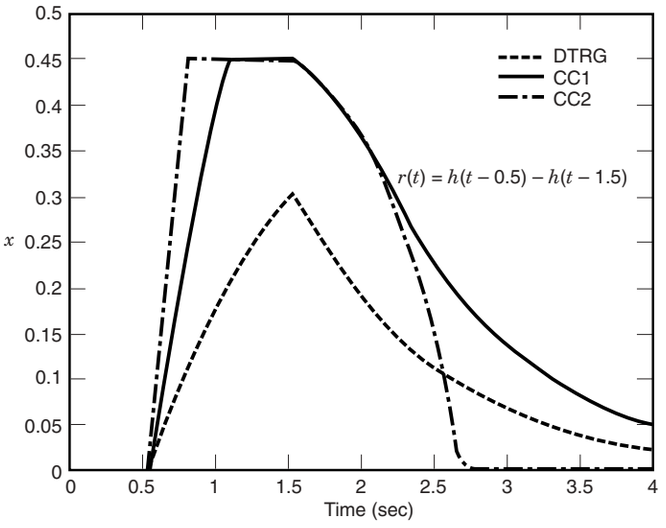

Figure 23. Comparison of responses to a statically inadmissible input.

Notice that while it is necessary to enforce $x \in X_I$ for the above control laws when the open-loop plant is unstable, neither Eq. (155) nor Eq. (159) requires that r' be statically admissible; that is, $|r'| > b/a$ is allowed when $x \in \text{int}(X_I)$. This is in contrast to Gilbert's DTRG, and the globally BIBO stable LQT controller constructed in Refs. 32 and 40, which restrict the feasible reference signal to statically admissible values. By relaxing the requirement that the modified reference signal be statically admissible, the tracking performance is enhanced. Simulation results are shown in Figs. 23 and 24. These concepts are applied to higher-order plants and flight control in Refs. 42, 43, and 49.

SUBOPTIMAL APPROACH

A single-input continuous-time control-constrained plant is considered:

$$\dot{x} = Ax + bu, \qquad -1 \le u \le 1$$
$$y = cx$$

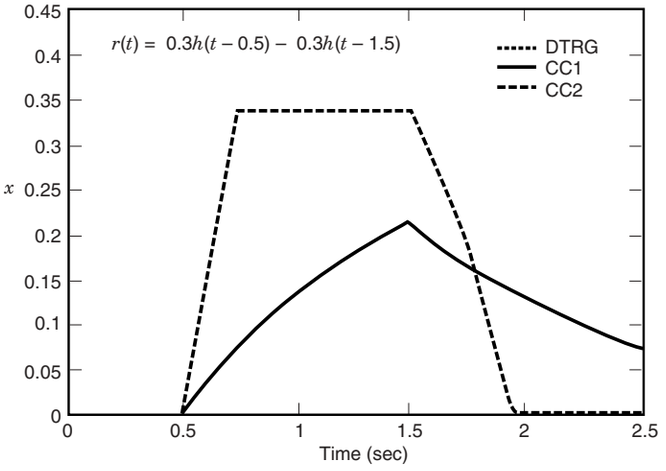

Figure 24. Comparison of responses to a statically admissible input.

The scalar controlled output variable y must track the exogeneous reference signal r. The point of departure is LQT. Hence, a tracking linear control law is specified; the simplest such control law is $u = k_x x + k_r r$, where $k_r = -1/cA_{cl}^{-1}b$ and $A_{cl} = A - bk_x$. Hence,

$$u = k_x x - \frac{1}{cA_{cl}^{-1}b} r$$

and the linear inner-loop is

$$\dot{x} = A_{cl}x - \frac{1}{cA_{cl}^{-1}b} br$$

The set of rest points of the closed-loop system is the one-dimensional subspace $x = (1/cA_{cl}^{-1}b)A_{cl}^{-1}br$, $-\infty < r < \infty$. Now, $-1 \le u \le 1$ yields the constraint on statically admissible r:

$$-1 \le \frac{k_x A_{cl}^{-1}b - 1}{cA_{cl}^{-1}b} r \le 1$$

Thus, the set of statically admissible reference signals is $-r_s \le r \le r_s$, where $r_s = cA_{cl}^{-1}b/(1 - k_x A_{cl}^{-1}b)$, and the set of admissible rest points is the segment

$$X_r = \{x \mid x = \frac{1}{cA_{cl}^{-1}b} A_{cl}^{-1}br, -r_s \le r \le r_s\}$$

BIBO stabilty enforcement requires the construction of certain invariant sets in the state space.

The maximal statically admissible set X_s is characterized as follows: The rest state which corresponds to the constant reference signal r is

$$x_r = \frac{1}{cA_{cl}^{-1}b} A_{cl}^{-1}br, \qquad -r_s \le r \le r_s$$

It is convenient to use the perturbation state $x : e; x - x_r$ and scale $r : e\ r/r_s$ whereupon the dynamics are transformed into

$$\dot{x} = A_{cl}x \tag{160}$$

and the saturation avoidance constraint is transformed into

$$-1 + r \le k_x x \le 1 + r \tag{161}$$

Now, fix r, $-r_s \le r \le r_s$, and determine the largest set contained in the slab in Eq. (161) which is invariant under Eq. (160); denote this maximal output admissible set by $X_s(r)$. Finally, the maximal statically admissible set X_s is

$$X_s = \cup_{-1 \le r \le 1} X_s(r)$$

Proposition. The set $X_s \subset R^n$ is bounded iff the pair (A, k_x) is observable.

Remark. The solution of well-posed inner-loop optimal control problems automatically renders the pair (A, k_x) observable. Hence, the set X_s will be bounded and BIBO stability can be enforced.

In general, the maximal statically admissible set X_s is compact and convex. However, X_s is not polyhedral; that is, it is

not finitely determined and, moreover, is rather difficult to construct.

In the two-dimensional case ($n = 2$), it is possible to obtain a closed-form representation of X_s: The boundaries of the r cross-sections $X_s(r)$ are trajectories of the unforced system; the boundaries of X_s are the envelopes of the above mentioned family of trajectories. For the control constrained system,

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 2 \end{bmatrix} u(t)$$
$$y(t) = [1 \quad 0]x(t)$$
$$-1 \le u(t) = [-9 \quad -2.5]x(t) + 8r(t) \le 1$$

X_s is shown in Fig. 25.

In higher dimensions, suboptimal (i.e., smaller statically admissible or invariant) sets, which are easy to characterize, are certainly useful. In engineering applications, guaranteeing BIBO stability for a reasonably large but bounded set of initial states which contains the origin is often sufficient. In this respect, the following holds.

Let P be the real symmetric positive definite solution of the Lyapunov equation

$$A_{cl}^{\mathrm{T}} P + P A_{cl} = -Q$$

where Q is a somewhat arbitrary, real, symmetric, and positive definite matrix. The ellipsoid

$$E_s = \left\{ x \,|\, x^T P x \le \frac{1}{k_x P^{-1} k_x^{\mathrm{T}}} \right\}$$

is statically admissible. Furthermore, $X_r \subset E_s$, provided that $(cA_{cl}^{-1}b)^2 > (k_x P^{-1} k_x^{\mathrm{T}})(b^{\mathrm{T}}(A_{cl}^{-1})^{\mathrm{T}}b)$.

It is now fairly easy to synthesize the outer-loop nonlinearity N, namely, the nonlinear control law $r' = r'(x, r)$. For example, consider the Control Strategy 1: r' is chosen to minimize $|r - r'|$, subject to the saturation avoidance constraint

$$-1 \le k_x x - \frac{1}{cA_{cl}^{-1}b} r' \le 1$$

provided that

$$x^{\mathrm{T}} P x < \frac{1}{k_x P^{-1} k_x^{\mathrm{T}}}$$

For x such that

$$x^{\mathrm{T}} P x = \frac{1}{k_x P^{-1} k_x^{\mathrm{T}}}$$

r' must satisfy the additional invariance-enforcing inequality

$$r' \frac{2}{cA_{cl}^{-1}b} b^{\mathrm{T}} P x \ge -x^{\mathrm{T}} Q x$$

Feasibility is guaranteed, by construction.

Hence, a dual-loop tracking controller as shown in Fig. 20 can be constructed. For all measurable exogeneous reference signals r and for all initial states in the ellipsoid E_s, good small-signal performance is achieved, high-amplitude signals are optimally tracked, and BIBO stability is guaranteed.

CONCLUSION

The basic concept of the stability of dynamical systems is carefully developed, and its application to the design of feed-

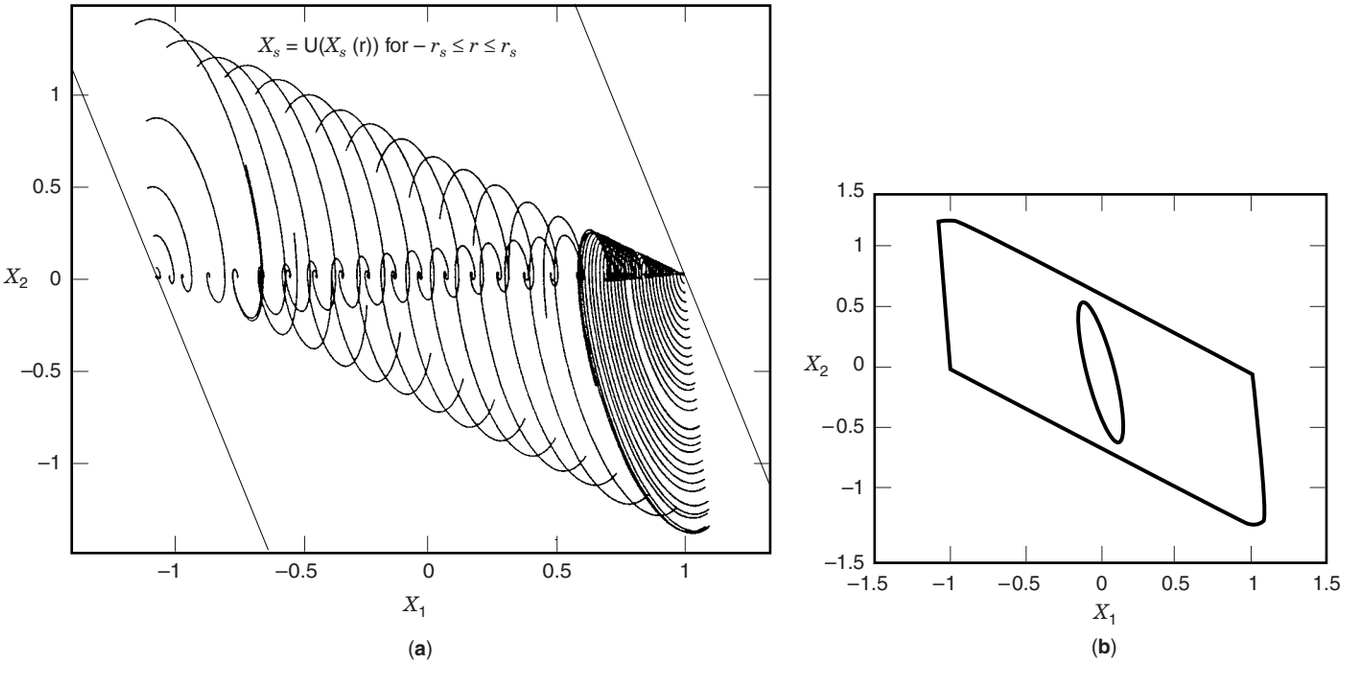

Figure 25. (a) Maximal statically admissible set. (b) Comparison of X_s and the suboptimal ellipse.

back control systems is presented. Classical stability theories for linear and nonlinear feedback control systems are discussed and the Lyapunov and Routh–Hurwitz stability criteria are developed. Linearization of nonlinear dynamics in terms of the Jacobian matrix is presented. Attention is given to conventional control design and to frequency-domain methods: The root locus and Nyquist stability criteria are developed and, in addition, the Bode plot and Nichols chart-based methods for stability, and degree of stability determination, are discussed. The emphasis then shifts to the stability of control systems comprised of linear plants and actuators which are subject to saturation. The state of the art with respect to nonlinear controller synthesis methods that address tracking control in the presence of actuator constraints is examined. The saturation avoidance approach is pursued. Special attention is given to tracking performance and to the BIBO stability of the closed-loop nonlinear control system. Results based on the concept of invariant sets—namely, maximal output admissible sets, maximal statically admissible sets, and the linear quadratic tracking (LQT) control method—are presented. These results are applied to the synthesis of a dual-loop nonlinear control structure. The on-line computational requirements are modest. Good tracking performance of high-amplitude command signals is achievable, small-signal performance is preserved, and BIBO stability is guaranteed—provided that the command signals are sufficiently smooth, and the initial state is contained in a predetermined invariant set.

BIBLIOGRAPHY

1. B. Porter, *Synthesis of Dynamical Systems,* London: Thomas Nelson, 1969.
2. W. J. Rugh, *Linear System Theory,* Englewood Cliffs, NJ: Prentice-Hall, 1993.
3. J. J. D'Azzo and C. H. Houpis, *Linear Control System Analysis and Design,* 3rd ed., New York: McGraw-Hill, 1988.
4. F. Csaki, *Modern Control Theories,* Budapest: Akademiai Kiado, 1972.
5. T. Kailath, *Linear Systems,* Englewood Cliffs, NJ: Prentice-Hall, 1980.
6. H. M. James, N. B. Nichols, and R. S. Phillips, *Theory of Servomechanisms,* Radiation Laboratory Series, Vol. 25, New York: McGraw-Hill, 1974.
7. J. E. Gibson, *Nonlinear Automatic Control,* New York: McGraw-Hill, 1963.
8. J. J. D'Azzo and C. H. Houpis, *Linear Control System Analysis and Design,* 4th ed., New York: McGraw-Hill, 1995.
9. K. N. Swamy, On Sylvester's criterion for positive semidefinite matrices, *IEEE Trans. Autom. Control,* **AC-18**: 306, 1973.
10. K. Ogata, *Modern Control Engineering,* Englewood Cliffs, NJ: Prentice-Hall, 1970.
11. N. Minorsky, *Theory of Nonlinear Systems,* New York: McGraw-Hill, 1969.
12. R. E. Kalman and J. E. Bertram, Control system analysis and design via the second method of Liapunov, *J. Basic Eng.,* **80**: 371–400, 1960.
13. E. A. Guillemin, *The Mathematics of Circuit Analysis,* New York: Wiley, 1949.
14. V. Singh, Comments on the Routh-Hurwitz Criterion, *IEEE Trans. Autom. Control,* **AC-26**: 612, 1981.
15. B. Porter, *Stability Criteria for Linear Dynamical Systems,* New York: Academic, 1968.
16. K. J. Khatwani, On Routh-Hurwitz criterion, *IEEE Trans. Autom. Control,* **AC-26**: 583–584, 1981.
17. S. K. Pillai, On the ϵ-method of the Routh–Hurwitz criterion, *IEEE Trans. Autom. Control,* **AC-26**: 584, 1981.
18. D. M. Etter, *Engineering Problem Solving with MATLAB,* Engelwood Cliffs, NJ: Prentice-Hall, 1993.
19. V. Krishnamurthi, Implications of Routh stability criteria, *IEEE Trans. Autom. Control,* **AC-25**: 554–555, 1980.
20. F. R. Gantmacher, *Applications of the Theory of Matrices,* New York: Wiley-Interscience, 1959.
21. *MATLAB Version 4,* Natic, MA: The MathWorks, Inc., 1993.
22. H. Nyquist, Regeneration Theory, *Bell Syst. Tech. J.,* **11**: 126–147, 1932.
23. H. W. Bode, *Network Analysis and Feedback Amplifier Design,* Princeton, NJ: Van Nostrand, 1945, Chapter 8.
24. T. F. Schubert, Jr., A quantitative comparison of three Bode straight-line phase approximations for second-order, underdamped systems, *IEEE Trans. Educ.,* **40**: 135–138, 1997.
25. A. A. Andronow, A. A. Vitt, and S. E. Chaikin, *Theory of Oscillations,* Reading, MA: Addison-Wesley, 1966.
26. R. S. Baheti, Simple anti-windup controllers, in *Proc. 1989 Amer. Control Conf.,* Vol. 2, 1989, pp. 1684–1686.
27. D. S. Bernstein, Optimal nonlinear, but continuous feedback control of systems with saturating actuators, in *Proc. 32rd Conf. Decision Control,* December 1993, pp. 2533–2537.
28. D. S. Bernstein, Actuator saturation and control system performance, in *Proc. 32rd Conf. Decision Control,* 1993, pp. 2420–2421.
29. P. J. Campo and M. Morari, Robust control of processes subject to saturation nonlinearities, *Comput. Chem. Eng.,* **14** (4/5): 343–358, 1990.
30. P. R. Chandler, M. J. Mears, and M. Pachter, A hybrid LQR/LP approach for addressing actuator saturation in feedback control, in *1994 Conf. Decision Control,* Dec. 14–16, 1994, pp. 3860–3867.
31. E. G. Gilbert and K. T. Tan, Linear systems with state and control constraints: The theory and application of maximal output admissible sets, *IEEE Trans. Autom. Control,* **AC-36**: 1008–1020, 1991.
32. E. G. Gilbert, I. Kolmanovsky, and K. T. Tan, Nonlinear control of discrete-time linear systems with state and control constraints: A reference governor with global convergence properties, in *Proc. 33rd Conf. Decision Control,* December 1994, pp. 144–149.
33. R. Hanus, M. Kinnaert, and J. L. Henrotte, Conditioning technique, a general anti-windup and bumpless transfer method, *Automatica,* 729–739, 1987.
34. R. A. Hess and S. A. Snell, Flight control design with actuator saturation: SISO systems with unstable plants, in *33rd Aerospace Sciences Meeting and Exhibit, AIAA,* paper No 95-0336, January 1995; also published in the *AIAA J. Guidance and Control,* **19** (1): 191–197, 1996, and see also the Errata in *AIAA J. Guidance and Control,* **19** (6): 1200.
35. I. Horowitz, Feedback systems with rate and amplitude limiting, *Int. J. Control,* **40**: 1215–1229, 1984.
36. P. Kapasouris, M. Athans, and G. Stein, Design of feedback control systems for unstable plants with saturating actuators, in *IFAC Symp. Nonlinear Control Syst. Design,* June 1989, pp. 302–307.
37. P. Kapasouris and M. Athans, Multivariable control systems with saturating actuators antireset windup strategies, in *Proc. 1985 Amer. Control Conf.,* 1985, pp. 1579–1584.
38. P. Kapasouris and M. Athans, Control systems with rate and magnitude saturation for neutrally stable uncertain plants, in *Proc. 29th Conf. Decision Control,* Vol. 6, 1990, pp. 3407–3409.

39. N. J. Krikelis and S. K. Barkas, Design of tracking systems subject to actuator saturation and integrator windup, *Int. J. Control,* **39**: 667–683, 1984.
40. R. B. Miller, A New Approach to Manual Tracking Flight Control with Amplitude and Rate Constrained Dynamic Actuators, Ph.D. dissertation, Air Force Institute of Technology, March 1997.
41. R. B. Miller and M. Pachter, Manual tracking control with amplitude and rate constrained actuators, in *Proc. 1996 Conf. Decision Control,* December 1996, pp. 3159–3164.
42. R. B. Miller and M. Pachter, Manual control with saturating actuators, in *Proc. 1997 Eur. Control Conf.,* July 1–4, 1997.
43. R. B. Miller and M. Pachter, Maneuvering flight control with actuator constraints, *AIAA J. Guidance, Control Dynamics,* **20** (4): 729–734, 1997.
44. C. Moussas and G. Bitsoris, Adaptive constrained control subject to both input-amplitude and input-velocity constraints, in *Proc. IEEE Int. Conf. Syst., Man Cybern.,* Vol. 4, 1993, pp. 601–606.
45. M. Pachter and D. H. Jacobson, Observability with a conic observation set, *IEEE Trans. Autom. Control,* **AC-24**: 632–633, 1979.
46. M. Pachter, P. R. Chandler, and M. Mears, Control reconfiguration with actuator rate saturation, in *Proc. Amer. Control Conf.,* June 1995, pp. 3495–3499.
47. M. Pachter, P. R. Chandler, and L. Smith, Velocity vector roll control, in *1996 AIAA Guidance, Navigation Control Conf.,* San Diego, CA, July 1996, AIAA paper No. 96-3867.
48. M. Pachter, P. R. Chandler, and L. Smith, Maneuvering flight control, in *1997 Amer. Control Conf.,* Albuquerque, NM, June 4–6, 1997.
49. M. Pachter and R. B. Miller, Manual flight control with saturating actuators, *IEEE Control Syst. Magazine,* **18**: 10–19, 1997.
50. V. M. Popov, *The Hyperstability of Control Systems,* New York: Springer, 1973.
51. A. A. Rodriguez and Y. Wang, Saturation Prevention Strategies for an Unstable Bank-To-Turn (BTT) Missile: Full Information, Ph.D. dissertation, Department of Electrical Engineering, Arizona State University, Tempe, AZ.
52. A. G. Tsirukis and M. Morari, Controller design with actuator constraints, in *Proc. Conf. Decision Control,* 1992, pp. 2623–2628.

J. J. D'AZZO
M. PACHTER
Air Force Institute of Technology
(AFIT/ENG)