CONTROL SYSTEM STABILITY

CHARACTERISTIC EQUATION: The overall transfer function for a feedback control system is: \( TF = G / [1+GH] \). The \( G \) and \( H \) functions can be put into the form:

\[
G(S) = \frac{A(S)}{B(S)} \quad H(S) = \frac{X(S)}{Y(S)}
\]

where \( A \), \( B \), \( X \), \( Y \) are polynomials. Substitution into the TF gives:

\[
TF = \frac{A}{B} / [1 + \frac{A}{B} \frac{X}{Y}] = \frac{AY}{BY + AX}.
\]

The transfer function can also be reduced to a ratio of two polynomials \( N(s) \) and \( D(s) \). In terms of these polynomials the characteristic equation is: \( D(S) = 0 \). Thus the characteristic equation in terms of \( A \), \( B \), \( X \), \( Y \) is: \( AX + BY = 0 \).

The \( GH \) function is: \( GH = A/B \frac{X}{Y} = \frac{AX}{BY} = N/D \). So the characteristic equation in terms of the \( GH \) function is:

\[
N + D = 0.
\]

Note that the characteristic equations for the subsytems are all contained in \( D(S)=0 \). Often \( D(S) \) is in factored form: so simple inspection tells if the subsystems are stable or unstable. This is not the case for the overall system because, even though both \( N(S) \) and \( D(S) \) may be in factored form, adding them destroys this.
ROOT LOCUS PLOTS: As some parameter of a system is varied, each root of its characteristic equation moves around in the S plane and traces out a path known as a Root Locus. The Root Locus Method is systematic set of sketching rules based on the GH function for finding approximate location of these paths. Numerical schemes for finding roots of polynomials can now be used to find Root Locus paths exactly. So the Root Locus Method is obsolete. However the paths themselves are very important because they show system parameter values corresponding to the onset of instability. Root Locus Plots for some simple systems are given in Figure 1. To generate each plot, the parameter K was varied from 0 to $\infty$.

ROUTH-HURWITZ CRITERIA: These criteria infer stability information directly from the coefficients in the characteristic equation. The method is based on the theorem of residues. It is rarely derived from first principles in controls text books. It shows that, when some of the coefficients of the characteristic equation are positive and some are negative, the system is unstable. It also shows that a zero coefficient implies that the best a system can be is borderline stable. As a bare minimum, for stable operation of a system, all of the coefficients must be nonzero and all must have the same sign. Consider the cubic characteristic equation:
\[ A S^3 + B S^2 + C S + D = 0 \]

where \( A \) is positive. For this case, Routh-Hurwitz shows that, for stable operation, all coefficients must be positive, and they must also produce a positive value when substituted into the test function \( X=BC-AD \). Consider the quartic characteristic equation:

\[ A S^4 + B S^3 + C S^2 + D S + E = 0 \]

where \( A \) is positive. For this case, Routh-Hurwitz shows that, for stable operation, all coefficients must be positive, and they must also produce positive values when substituted into the test functions \( X=BC-AD \) and \( Y=DX-B^2E \). Consider the quintic equation:

\[ A S^5 + B S^4 + C S^3 + D S^2 + E S + F = 0 \]

where \( A \) is positive. For this case, Routh-Hurwitz shows that, for stable operation, all coefficients must be positive, and they must also produce positive values when substituted into the test functions \( X=BC-AD \) \( Z=BE-AF \) and \( Y=(DX-BZ)Z-X^2F \).

NYQUIST : A Nyquist Plot is a closed contour in the GH plane (or the 1+GH plane). It is obtained by mapping a closed contour in the S plane to the GH plane (or the 1+GH plane) using GH (or 1+GH) as a mapping function. The closed contour in the S plane surrounds the entire right half or unstable half of the S plane. A typical
mapping is shown in Figure 2. Stability is inferred from the plot in the GH plane (or the 1+GH plane). Development of the Nyquist Concept is based on the 1+GH function:

\[
1 + GH = 1 + \frac{N}{D} = \frac{N + D}{D}.
\]

The overall characteristic function is \( N+D \): the subsystems characteristic function is \( D \). The roots of \( N+D=0 \) are called the zeros of the 1+GH function while the roots of \( D=0 \) are called the poles of the 1+GH function. Zeros are roots of the overall characteristic equation while poles are roots of the subsystem characteristic equations. At a zero \(|1+GH|=0\) while at a pole \(|1+GH|=\infty\). One can construct a 3D image of \(|1+GH|\) by taking the S plane as a horizontal plane and plotting \(|1+GH|\) vertically. At a zero the image would touch the S plane. At a pole its height above the S plane would be infinite. The plot could be used to determine the stability of the system and its subsystems.

One could factor 1+GH to get its zeros \( Z \) and poles \( P \):

\[
1 + GH = \frac{K (S-Z_1) (S-Z_2) \cdots (S-Z_n)}{(S-P_1) (S-P_2) \cdots (S-P_m)}.
\]

In the S plane, each \((S-Z)\) or \((S-P)\) factor is basically a vector with radius \( r \) and angle \( \theta \): \( r\angle\theta \). A typical vector is shown in Figure 3. What happens to these vectors as the tip of the S vector moves once in a clockwise sense around the contour which surrounds the entire right half of the S plane? As shown in Figure 4,
vectors inside rotate clockwise 360° while vectors outside only nod up and down. What are the implications of this for the 1+GH function? Consider the Polar Form of 1+GH:

\[ K \left( \prod r_z \angle \sum \theta_z \right) / \left( \prod r_p \angle \sum \theta_p \right) = R \angle \Theta \]

where \( \Pi \) indicates product and \( \Sigma \) indicates sum. Zeros inside cause clockwise rotations of 1+GH: poles inside cause counterclockwise rotations of 1+GH. Only zeros and poles inside cause such rotations: zeros and poles outside only cause 1+GH to nod up and down. If clockwise rotations are considered positive and counterclockwise rotations are considered negative, then the net clockwise rotations of 1+GH must be: \( N = N_z - N_p \) where \( N_z \) is the number of zeros in the unstable half of the S plane while \( N_p \) is the number of poles there. For stable operation, \( N_z \) must be zero. When \( N_z \) is positive, the system is unstable. Inspection of D gives \( N_p \). Inspection of the 1+GH plot gives \( N \). Substitution into \( N_z = N + N_p \) gives \( N_z \). When a vector is drawn from the origin of the 1+GH plane to the 1+GH plot, \( N \) is the net number of times that this vector rotates clockwise when its tip moves along the plot.

The minus one point on the real axis in the GH plane corresponds to the origin in the 1+GH plane. This implies that a rotation of the GH vector drawn from the minus one point in the GH plane is equivalent to a rotation of the 1+GH vector drawn from the origin in the 1+GH plane. So one can get \( N \) from inspection of the GH plot or the 1+GH plot. This is illustrated in Figure 5.
The basic Nyquist contour in the S plane consists of the imaginary axis and an infinite radius semicircle. This contour surrounds the entire right half or unstable half of the S plane. Sometimes there are poles of GH on the imaginary axis in the S plane. They are usually located at the origin. At a pole GH is infinite. To avoid this, the contour is indented locally with an infinitesimal radius counterclockwise semicircle centered on the pole.

To construct a GH plot, each section of the Nyquist contour is mapped separately. The infinite radius semicircle usually maps to the origin in the GH plane. An infinitesimal radius semicircle always maps to an infinite radius semicircle in the GH plane. Each pole on the imaginary axis produces one semicircle in the GH plane. The imaginary axis in the S plane can be mapped point by point to the GH plane. The negative imaginary axis portion is a mirror image of the positive imaginary axis portion. The location of these portions relative to the minus one point is usually critical. One can get a rough sketch of these portions by first fixing the small and large \( \omega \) end points. One then examines the GH function to see if it is possible to make it purely real or purely imaginary. Purely real means there is a real axis crossover while purely imaginary means there is an imaginary axis crossover. With known end points and crossovers, one can quickly sketch the plot.
Root Locus Plots

\[ G(s)H(s) = \frac{K(s+a)}{(s+\alpha+j\beta)(s+\alpha-j\beta)} \]

1

\[ G(s)H(s) = \frac{K(s+a)}{s(s+b)} \]
$S = A + Bj$

$F = a + bj$

$S - F = (A - a) + (B - b)j = r/\theta$