APPENDIX E

FEEDBACK CONTROL SYSTEM CONCEPTS $^{25,30}$

The term "Feedback Control System" is applied to a control system which compares a quantity to be controlled with a reference or desired value and operates on the error between these to bring the controlled quantity towards the desired value.

The closeness with which the controlled quantity is brought towards the desired value is a function of the type of control system. Regulators or type "O" servomechanisms require some finite error in the steady state. Type 1 servomechanism or controllers with "reset" (in process control terminology) have integration in the controller and wipe out the error to zero in the steady state.

One of the greatest aids in understanding control systems is the block diagram. This was introduced in Appendix C. A very common configuration of a feedback control is shown on Fig. E-1.

\[ \text{FIGURE E-1} \]

Here $G_p$ is the process transfer function
$G_c$ is the controller forward function
$H_c$ is the feedback function
$R$ is the reference quantity or set point
$C$ is the controlled variable
$m$ is the input to the process manipulated by the controls
$u$ is the load or disturbance to the process.

This configuration may be understood better through an example. Take the case of a level control system:

$G(p)$ could be the transfer function $K/s$ of a tank, $c$ being the level of resident fluid,
$m$ the input controlled, flow and
$u$ being the uncontrolled (disturbance) drawdown from the tank
$R$ is the reference level, or set point
$H$ could be the sensor transfer function, could be a simple lag \( \frac{1}{1+5T} \),
$G_c$ is the controller function which could be a simple proportional gain $K_p$ or some more complex function.

Using the expressions on block diagram reduction of Appendix C, we can write down

\[
C = \frac{G_c G_p}{1 + G_c G_p H} R + \frac{G_p}{1 + G_c G_p H} u
\]  

(E-1)

The effect of closing the loop can be appreciated from equation E-1.
For instance the effect of a disturbance \( u \) on the system output \( C \) would have been \( G_p u \) with the loop open. This has been reduced by a factor \( \frac{1}{1+G_c G_p H} \) by closing the control loop. If for instance \( G_c G_p H \) had a steady state value of 9, the effect of the disturbance will be reduced 10 to 1. If \( G_c G_p H \) were to have an integration, (equivalent to having infinite gain in the steady state) the effect of the disturbance would have been zero in the steady state.

Another important point which may be drawn from equation E-1 is the expression for \( C/R \), i.e., the change of the controlled variable in response to a change in desired or reference value.

\[
\frac{C}{R} = \frac{G_c G_p}{1+G_c G_p H} \quad (E-2)
\]

For very large \( G_c G_p \), \( C/R = \frac{1}{H} \). This expression shows the importance of accuracy and linearity in the feedback element \( H \), since for systems with large values of \( G_c G_p H \) the controlled variable \( C \) depends mainly on the feedback element \( H \), and the value of reference \( R \).

Feedback control does not only reduce the effect of disturbances and drifts in the steady state. It also causes the output to respond more rapidly to the command of the references or to counteract dynamically the effects of load disturbances. Here we get involved with the important subject of response and stability of closed loop systems.

Closed Loop and Open Loop Time Response of Simple Systems

The concept of response and stability will be introduced with a couple of simple examples. Take the case of an open circuit generator whose terminal voltage response to a change in field volts is governed by a single time constant as described in Fig. E-2.

\[
\Delta E_{fd} \quad \frac{1}{1 + ST_{do}'} \quad \Delta e_t
\]

\[
\Delta E_{fd} \quad \frac{\Delta e_t}{t \rightarrow T'_{do}}
\]

**FIGURE E-2**

Let us now regulate terminal voltage by means of an idealized regulator and exciter which may be represented by a simple gain as in Fig. E-3.
The response of the closed loop

\[
\frac{\Delta e_t}{\Delta V_{\text{Ref}}} = G = \frac{K}{1 + sT_{\text{do}}} \cdot \frac{1 + sT_{\text{do}}}{1 + \frac{K}{1 + sT_{\text{do}}}}
\]

\[= \frac{K}{1 + K + sT_{\text{do}}} \cdot \frac{K}{1 + K (1 + \frac{sT_{\text{do}}}{1 + K})} \]

(E-3)

For large \( K \) we note that equation E-3 has a steady state gain of almost unity but a response time constant of \( T_{\text{do}}/(1+K) \) which is \((1+K)\) times faster than the open loop response of Fig. E-2.

Fig. E-4 shows the comparison of the open and closed loop performances.

Note that the response of the closed loop is considerably faster than that of the open loop. This is due to the forcing action of the field in response to the high gain operating on the error.

On an idealized system containing only one time constant (first order system) such as in Fig. E-3 there is theoretically no limit to the value of the gain \( K \) that can be applied. In practice most systems exhibit more than one time constant, i.e.; are of higher than first order, and a limit
to the gain of the closed loop and hence a limit to closed loop's speed of response is reached due to stability considerations.

Let us illustrate by assuming the regulator exciter to be described by one time constant as in Fig. E-5.

![Block Diagram](image)

FIGURE E-5

Again solving for $\Delta \eta / \Delta V_{ref}$ with the use of the $G/(1+GH)$ formula

\[ \frac{\Delta \eta}{\Delta V_{ref}} = \frac{K}{(1+K)} \left[ \frac{1}{1 + s\frac{(T_c + T_d)}{(1+K)} + s^2 \frac{T_c T_d}{(1+K)}} \right] \]  \hspace{1cm} (E-4)

The nature of the closed loop response can be derived from the roots of the denominator of Equation E-4, which are the roots of the closed loop characteristic equation

\[ 1 + GH = 0 \]

Depending on the value of $K$, the roots of equation E-4 can be real or complex, the higher the value of $K$ the more oscillatory (less damped) are the roots.

The quadratic form of the denominator can be expressed as

\[ 1 + \frac{2\zeta}{\omega_0} s + \frac{s^2}{\omega_0^2} \]  \hspace{1cm} (E-5)

where

\[ \frac{2\zeta}{\omega_0} = \frac{T_c + T_d}{(1+K)} \]  \hspace{1cm} (E-6)

and

\[ \frac{1}{\omega_0^2} = \frac{T_c T_d}{(1+K)} \]  \hspace{1cm} (E-7)

The roots of equation E-5 are

\[ s_1, s_2 = \left[ -\zeta \pm j \sqrt{1 - \zeta^2} \right] \omega_0 \]  \hspace{1cm} (E-8)
where $\omega_0$ is known as the natural frequency of oscillation and $\zeta$ is the damping ratio. A damping ratio $\zeta = 1$, known as critical damping, yields two equal real roots $s_1, s_2 = -\omega_0$. Damping ratios less than 1 yield complex roots while those greater than one yield real (non-oscillatory) roots.

In terms of the parameters $K, T_e$ and $T_{do}'$, equation E-6 and equation E-7 yield

$$\zeta = \frac{T_e + T_{do}'}{2 \sqrt{1 + K} \left( T_e T_{do}' \right)} \quad (E-9)$$

$$\omega_0 = \sqrt{\frac{1 + K}{T_e T_{do}'}} \quad (E-10)$$

From equation E-9 and equation E-10 we note that the higher $K$, the higher the natural frequency and the lower the damping ratio. The second order system of Fig. E-5 cannot become unstable, i.e., cannot have roots with negative values of $\zeta$. However, it can approach exhibiting sustained oscillations as $\zeta \to 0$ which is unacceptable performance. Systems with characteristic equations of higher orders can easily exhibit instability with increasing loop gains.

The inverse Laplace transform of expression E-4 multiplied by $1/S$ (for the input step) yields the time function

$$\Delta t = \frac{K}{(1+K)} \left[ 1 + \frac{\omega_0 t}{\sqrt{1-\zeta^2}} \sin \left( \sqrt{1-\zeta^2} \omega_0 t - \tan^{-1} \frac{1-\zeta^2}{\zeta} \right) \right] \quad (E-11)$$

Figure E-6 shows the form of the response expression $E-11$ as function of normalized time "$\omega_0 t$" for various values of $\zeta$. 
STABILITY OF CLOSED LOOP SYSTEMS

From the above examples it becomes quite clear that the stability of closed loop systems can be investigated from knowledge of the roots of the characteristic equation

\[ 1 + GH = 0 \]

Several methods are available to determine whether or not some of the roots of the characteristic equation lie in the right hand plane (a condition signifying instability). Such methods as Routh's criterion, root locus, etc. have their special application and it is not the intent here to expound further on these methods which may be readily found in the literature. We will merely explore briefly some of the very widely used Frequency Response techniques and Nyquist stability criteria. The table below summarizes briefly the features and names of some of the techniques used for determining stability.

### STABILITY CRITERIA CHARACTERISTICS

<table>
<thead>
<tr>
<th>Method</th>
<th>Answer Obtained</th>
<th>Information Required</th>
<th>Application</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routh</td>
<td></td>
<td>Closed-loop characteristic polynomial</td>
<td>Analysis</td>
<td>Difficult to assess effect of parameter variations</td>
</tr>
<tr>
<td>Hurwitz</td>
<td></td>
<td></td>
<td></td>
<td>Involved computation required</td>
</tr>
<tr>
<td>Meirov</td>
<td></td>
<td></td>
<td></td>
<td>Only computation needed is long division</td>
</tr>
<tr>
<td>Wall</td>
<td>Yes-or-no stability</td>
<td>Closed-loop characteristic polynomial</td>
<td>Analysis</td>
<td>Applicable to sampled data systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schur</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root Locus</td>
<td>Complete system response</td>
<td>Open-loop transfer function, factored form</td>
<td>Analysis and synthesis</td>
<td>Can be extended to time delay systems</td>
</tr>
<tr>
<td>Nyquist</td>
<td>Stability and approximate time response</td>
<td>Open-loop transfer function measured or all frequencies</td>
<td></td>
<td>Application to time delay systems</td>
</tr>
<tr>
<td>Bode</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dzung</td>
<td>Yes-or-no stability</td>
<td>Closed-loop characteristic polynomial</td>
<td>Analysis</td>
<td></td>
</tr>
<tr>
<td>Mickallov</td>
<td>Absolute and relative stability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leonhard</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Frequency Response

The concept of the operational impedance and the impedance to a sinusoidally varying excitation function was developed in example 2 of Appendix A.

A transfer function is an operational expression much like that of impedance. The complex number obtained by substitution of $s = j\omega$ in the transfer function gives information on the steady state sinusoidal response of the output of the function to a sinusoidal input excitation of frequency $\omega$. The absolute value of this number corresponds to the magnitude ratio of the output sinusoid to the input sinusoid while the phase angle of the complex number expressed in polar coordinates indicates the angle by which the output sinusoid leads or lags the input sinusoid.

Frequency response techniques use the magnitude and phase characteristics of transfer functions or combinations of transfer functions to derive a great deal of information about the stability and response performance of control systems.

The frequency response characteristics of some common transfer functions encountered in control systems are described in Fig. E-7.

Naturally the frequency response of a combination of transfer functions in series is easily obtained by taking the product of these functions, i.e., the magnitude is the product of the magnitudes of the individual functions and the overall phase angle is equal to the summation of phase angles of the individual transfer functions.

Nyquist Stability Criterion

Recall that the stability of a closed loop system was determined from properties of the characteristic equation $1 + GH = 0$.

The Nyquist criterion is a means used to determine whether or not $1 + GH$ has roots in the right hand half of the s plane. A rigorous derivation of the Nyquist criterion involves use of complex number theorems by Cauchey and examination of the number of cycles of phase rotation of the function $GH(j\omega)$ as $j\omega$ is taken around a closed path from $-j\omega$ to $+j\omega$.

Except for very unusual circumstances Nyquist's criterion applied to practical cases amounts to the following:

For a closed loop whose characteristic equation is $1 + GH$, the stability of the system can be derived by examining the frequency response characteristic of the open loop function $GH$. This is done by finding the phase angle of $GH$ at the frequency for which the magnitude of $GH$ is 1.0. If the phase angle is $180^\circ$ the system is borderline unstable. If the phase angle is more than $180^\circ$ lagging, the system is unstable.

The phase angle of $GH$ at the point where the magnitude of $GH$ is 1 is known as the phase angle at crossover and the frequency $\omega_c$ at this point is called the crossover frequency. Many guide rules have been established to relate the shape of the open loop frequency response function to the performance of the closed loop function. One point to remember is that the phase angle at crossover should in general not exceed 130 to 140°. For such cases the closed loop system response will be oscillatory with good damping, the output exhibiting an overshoot of about 25%. The frequency of oscillation of the closed loop is related to and closely approximates the crossover frequency. For smaller phase angles at crossover, in the order of $100^\circ$ the system in general looks critically damped.

Figure E-8 shows complex plots of typical $G(j\omega)H(j\omega)$ (open loop) functions as $\omega$ varies from zero to infinity. Such plots are called Nyquist diagrams. The third locus in Fig. E-8 shows the case of a conditionally stable system - one where either an increase or a decrease in loop gain
can cause instability. This is in contrast with the usual case where instability is only reached with increasing gain.

Figure E-9 shows typical time responses of the closed loop $\Delta C = \frac{G}{1+GH} \Delta R$ to a step change in reference $\Delta R$. 
Bode Theorems

From the foregoing discussion, we note that the evaluation of the overall open loop frequency response characteristic \( GH(j\omega) \) requires multiplication of transfer functions. The Bode method gives a simple technique for plotting transfer functions in terms of a log of magnitude plot and a phase angle plot.

If two transfer functions are written in polar form

\[
A = A_1 e^{j\theta_1}
\]

and

\[
B = B_1 e^{j\theta_2}
\]

Then

\[
AB = A_1 B_1 e^{j(\theta_1 + \theta_2)}
\]

That is, to obtain the frequency response of a product of transfer functions, the individual transfer function phases are added and their magnitudes are multiplied. If the magnitudes are expressed as \( \log A_1 \) and \( \log B_1 \), then \( \log (A_1 B_1) = \log A_1 + \log B_1 \). The Bode diagram plots the logarithm of magnitude and phase angle of the open loop function \( GH(j\omega) \) as separate functions of frequency.

The Bode plotting technique is based on asymptotic characteristics of transfer functions expressed in factored form. Let us illustrate with examples:

Take the transfer function of a single lag time constant

\[
G(s) = \frac{K}{1 + sT}
\]

The frequency response of \( G(s) = G(j\omega) \)

\[
= \frac{K}{1 + j\omega T} = \frac{K}{1 + j(\frac{\omega T}{\omega_0})}
\]
where \( \omega_0 = \frac{1}{T} \)

The asymptotes of equation E-15 as \( \omega/\omega_0 \to 0 \) and as \( \omega/\omega_0 \gg 1 \) are

\[
K \text{ and } -j \frac{K}{\omega_0} \tag{E-16}
\]

The magnitude of \( G(j\omega) \) plotted in log scale versus \( \omega \) also on a log scale is shown on Fig. E-10 as are also the two asymptotes for \( \omega/\omega_0 \to 0 \) and \( \omega/\omega_0 \gg 1 \).

Figure E-10 also shows the plot of phase angle of equation E-15 as function of \( \omega_0 \). Note that plotted in this form all that is required is the location of the break frequency \( \omega_0 \) and the shape of the frequency response function is then quickly determined.

A lead function is likewise described by asymptotic straight line approximations of log magnitude versus log \( \omega \) as also shown on Fig. E-10. The actual function can be quickly determined with the use of templates which give appropriate corrections to the asymptotic straight line approximations a function of the normalized value of \( \omega/\omega_0 \).

Likewise, templates give the phase angle contribution of each lag or lead factor, i.e.; the angle contributed by a pole or zero. Obviously the phase angle of equation E-15 as \( \omega/\omega_0 \to 0 \) is \( 0^\circ \) and it is \( 90^\circ \) as \( \omega/\omega_0 \to \infty \). At \( \omega/\omega_0 = 1 \) it is \( 45^\circ \).

Figure E-11 shows typical nomograms for use with the Bode technique.

We will illustrate the use of Bode diagrams in the following example:

Consider the position control system described by the block diagram of Fig. E-12. Find the maximum gain \( K \) for which this positioning system will be stable, and also the value of gain \( K \) for which the phase angle at crossover is \( -135^\circ \), i.e.; the phase margin is \( 45^\circ \).

Without ever knowing anything about Bode methods, this problem can be solved easily by plotting phase angle and magnitude of "G" with \( K = 1.0 \) as function frequency \( \omega \). From these plots we can determine the frequencies for which the phase angle is \( 180^\circ \) and \( 135^\circ \) respectively. Determine also the magnitude \( |G| \) at these frequencies. Let these magnitudes be \( M_{180^\circ} \) and \( M_{135^\circ} \) respectively.

Then the values of gain \( K \) required to cause crossover at a phase angle of \( 180^\circ \) and \( 135^\circ \) respectively are

\[
K_{180^\circ} = \frac{1}{M_{180^\circ}} = 12
\]

and

\[
K_{135^\circ} = \frac{1}{M_{135^\circ}} = 1.92
\]

Figure E-13 contains a Bode plot of \( \frac{1}{s(1 + T_1s)(1 + T_2s)} \) by use of the asymptotic approximation technique. Phase angle values are marked along the curve from which the value of \( M_{135^\circ} \) and \( M_{180^\circ} \) can be obtained by interpolation.
\[ |M_2|\theta_2 = 0.2 \left(1 + \frac{\omega}{\omega_0}\right) \]

\[ \frac{1}{1 + j\frac{\omega}{\omega_0}} = |M_1|\theta_1 \]

\[ \begin{array}{c}
0^\circ \\
-20^\circ \\
-40^\circ \\
-60^\circ \\
-80^\circ \\
-100^\circ \\
\hline
0.1 & 0.2 & 0.4 & 0.6 & 1.0 & 2.0 & 4.0 & 6.0 & 8.0 & 10 & 20
\end{array} \]

\[ \begin{array}{c}
\theta_1 \\
\theta_2 \\
\theta_4 \\
\hline
\omega/\omega_0
\end{array} \]

FIGURE E-10
- Place arrow at frequency at which phase is desired.
- At each break frequency record angle on scale. Denominator breaks angle is negative. Numerator breaks angle is positive.
- Add angles algebraically to obtain phase of total function.

**Gain Correction**

- Place arrow at frequency at which gain is desired.
- Read number on gain correction scale opposite each break.
- Add numbers for all breaks. Note sign of up-breaks is opposite sign of correction numbers for down breaks.
- Resulting number of units on scale C is gain correction.
WHERE $T_1 = 0.1$ SEC

$T_2 = 0.5$ SEC

FIGURE E-12
\[ G(s) = \frac{1}{S(1 + S0.5)(1 + S0.1)} \]

**Single Slope Due to \( V_s \)**

**Double Slope**

**Triple Slope**

**Asymptotes**

**Actual Magnitude**

**Breaks**

\[ M_{135^\circ} = 0.52 \]
\[ M_{180^\circ} = 0.83 \]

**Rads / Sec**

\[ \omega_c^{135^\circ} \]
\[ \omega_c^{180^\circ} \]

**Figure E-13**
For the value of \( K \) which produces borderline stability, the phase angle is \( 180^\circ \) and the crossover frequency is \( \omega_c = 4.5 \text{ rads/sec} \). The crossover frequency for the case where phase margin is \( 45^\circ \) is \( \omega_c = 1.5 \text{ rads/sec} \).

Control design by frequency response techniques is concerned with the shaping of the frequency response of the controller function to provide the proper phase margin at crossover. Lead/lag functions of the form \((1 + T_1 s)/(1 + T_2 s)\) with \( T_1 > T_2 \) and lag lead functions with \( T_4 < T_2 \) are often used, with the location of \( T_1 \) and \( T_2 \) selected so as to produce the desired effect.

It is apparent that a high or infinite gain at zero frequency is necessary for good steady state and low frequency performance (low, or zero steady state error). The response of the control system is related to the crossover frequency, the higher this frequency the faster the control response, or bandwidth. However, a limiting constraint is the requirement of stability.

The rules of plotting Bode diagrams can be readily derived by elementary reasoning. These can be formulated as follows:

1. Determine the gain and slope at zero frequency.
   a) If all factors contain time constants \((1 + sT)\) both in the numerator as well as the denominator, then the initial gain, at \( \omega = 0 \), is equal to the steady state gain and the initial slope is zero.
   b) If \((s)^n\) appears in the denominator, the initial gain or magnitude of the function is infinite and the slope is such that the gain is decreased by a factor \( 10^n \) for every tenfold increase in frequency. The line can be located by making it pass through the point determined by \( \omega = 1 \) on the abscissa and overall gain on the abscissa.
   c) If \((s)^n\) appears in the numerator, the initial magnitude ratio is zero and the slope is such that the gain is increased by a factor of \( 10^n \) when the frequency is increased tenfold.

2. The initial line is carried to the first break frequency \( \omega_0 = 1/T_1 \) where \( T_1 \) is the longest time constant in the numerator or denominator.

3. At the first break, the change in slope is determined. Since the factor containing this longest time constant will be in the form \((1 + sT)^n\), the gain will change by a factor \( 10^n \) when the frequency changes 10 times.
   a) For instance, if the initial slope is zero and \( n = 2 \) (factor containing the longest time constant in the numerator of the transfer function), then the slope after the break will be such that the gain is increased a hundredfold \( (10^2) \) when the frequency changes by a decade. This is called a double break upward.
   b) As another example assume that the initial slope is \( S^2 \); i.e.; the gain increases by ten times per decade increase in frequency and that the longest time constant occurs in the denominator at a break frequency \( 1/T_1 \). Then the slope after the break will be horizontal since the downward slope contributed by the denominator time constant cancels the upward initial slope.

4. The slope thus determined after the first break is continued until the next break which is determined by the next longest time constant in numerator or denominator. The change in slope is determined as before and the process is repeated until all time constants have been accounted for.

Bode theorems relate the slope of the magnitude function to the phase angle. In general for minimum phase functions the phase angle can be approximately determined by the slope of the function.
A single slope, i.e.; magnitude decreasing 10 times for a tenfold increase in frequency carries about 90° phase lag. A slope of two (two decades per decade), i.e.; 100 times decrease in magnitude for every tenfold increase in frequency represents about 180° phase lag and so on.

**Frequency Response of the Closed Loop**

The closed loop function

\[ \frac{G}{1 + GH} \quad (E-16) \]

can also be plotted in terms of its gain and phase as function of frequency.

Some easy guide rules can again be used to approximate the shape of the gain versus frequency curve.

Equation E-16 can be approximated under two extreme conditions, i.e.; when \( GH >> 1 \) and when \( GH << 1 \). Under the first condition equation E-16 becomes nearly \( \frac{1}{H} \) and under the second condition equation E-16 is approximately \( G \). This leads to the following set of rules to obtain the approximate shape of the closed loop response.

1. Plot gain curves for \( G, \frac{1}{H} \) and \( GH \).
2. Follow \( G \) when \( GH < 1 \), i.e.; follow \( G \) if \( \frac{1}{H} > G \).
3. Change from the \( G \) curve to the \( \frac{1}{H} \) curve when \( GH > 1 \), i.e.; follow \( \frac{1}{H} \) if \( \frac{1}{H} < G \).
4. The amount of resonant humping at near the transition from one curve to another is a function of the phase angle of \( GH \) at crossover, i.e.; at the point where \( GH = 1.0 \). Evidently if the phase angle at this point is 180°, the resonant peak would reach infinity. The more oscillatory the system the greater the peak of the closed loop function. Figure E-14 illustrates the various points made above using the example of Fig. E-12 for an arbitrary gain \( K = 1.5 \).
\[ G = \frac{1.5}{S (1+S0.5)(1+S0.1)} \]

\[ H = 1.0 \]

FIGURE E-14