

# Chapter 2

## Elements of Control Theory

### 2.1 Introduction

Proper analysis of the myriad manufacturing processes requires a bewildering breadth of physical understanding, easily beyond the reach of any one specialist. However, the control approach taken in this text is intended to overcome this vastness by referring all processes to a common body of theory, that of feedback control systems. While this theory itself has many different aspects and many limitations, our objective is to develop a common perspective for viewing processes and their control based on linear system theory and classic control concepts. No direct background in control theory is assumed, only a familiarity with ordinary differential equations.

Throughout this chapter, we concentrate on a single example system: the DC motor servo control problem. This is not only a convenient system to model and analyze, it also is the basic building block for many manufacturing control systems. As a result, the material is important even for those with a solid grounding in linear control theory.

### 2.2 Causality/control model

The basic "control perspective" is one of an input-output system with clear causality. This means that regardless of the details of the particular process we plan to control, we can (and indeed must) describe the process in a causal or input-output fashion. As will be shown below, linear control theory admits only one type of process model (the linear, time invariant dynamic system) and the art of process modelling for control is to capture the essential aspects of the process in this form. (More discussion of this type of *control modeling* will be covered in Chapter 3.) Few if any of manufacturing processes can be considered simple linear systems, but in some cases the restrictions can be violated without negating the utility of control theory. However, while the power of the control approach is to provide a well defined model form, learning the inherent limitations of this form will be an important element in this text.

## 2.3 The Feedback Control system: Structure and Properties

The basic layout of a feedback or closed-loop control system is shown in Fig. 2.1. The essential elements of this system are shown in the figure, and include the **plant** (the system to be controlled), the **measurement** (sensors, estimators and signal conditioners), the **controller** (the system we create to implement a control law or algorithm) and the **error** junction (where the desired system outputs and the measured or estimated outputs are compared to generate the error). Also evident from this diagram are the various signals in the system. (Careful adherence to the definition of these signals is important; it will prevent confusion in later discussions when manipulating process control system gets more complex and less similar in form to Fig. 2.1).

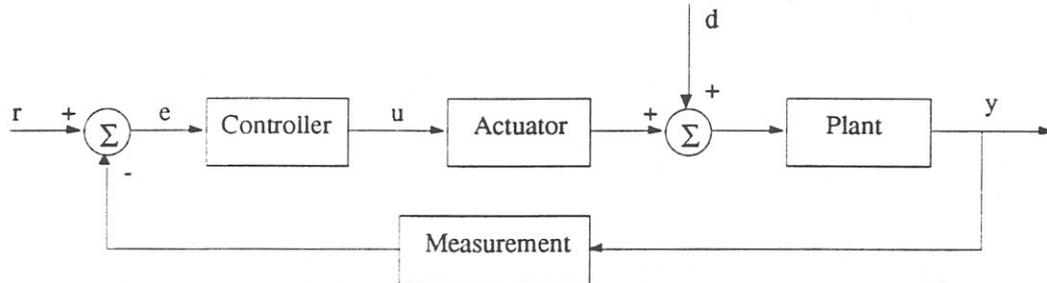


Figure 2.1 A Canonical Feedback Control System

The system **output**  $y$  is the physical quantity we wish to regulate. The input to the control system is actually the **desired output** or **setpoint**  $r$ . When measured outputs and setpoint are compared, we generate the **error**  $e$ , which is the signal upon which all of our control actions are based. The controller then takes this error and creates a **control output**  $u$ , which serves as an input (through an actuator) to the plant. Finally, there is the **disturbance**  $d$ , which we will define as any influence that causes undesired changes in the output. (This disturbance can enter the system in many places, not just that shown in Fig. 2.1)

### *Study Problem: The Position Servo*

To review the basics of classic control theory, we will deal with a single example system: the position servomechanism. Most manufacturing equipment, if it entails any level of automation, comprises, at least, a position servo. Combined with a digital computer, the position servo imbues any machine with programmable motion of great precision.

The problems in servo analysis and design are quite similar to those encountered in other “parameter” or machine control problems (such as force, pressure, flow, or temperature control) and is a system that is well modeled by standard linear system methods. However, as will be discussed in Chapter 3, linearity makes the specifics of servo controller design quite different from those of a typical manufacturing process. Accordingly, it is important to realize the idealized nature of the ensuing development, and that it is providing the basic underpinnings, but not the direct analog for manufacturing process control.

A simple position servo is shown in Fig. 2.2, and below it is a block diagram, showing how each element is connected. Notice that the system involves feedback, and that the key elements in a feedback system: measurement, comparison, error generation and actuation, are all included. In operation, the error signal is generated by a difference between the desired position and the actual position. This error is sent to the controller, where it may be amplified or filtered to create the actuator signal  $u$ . The actuator signal is sent to a power amplifier, where the signal is converted to a level of power sufficient to drive the motor. The motor torque causes the load to move, and this motion leads to a position change, which (if the system is correctly constructed) will reduce the error to zero. Thus the system automatically causes the input and output to match.

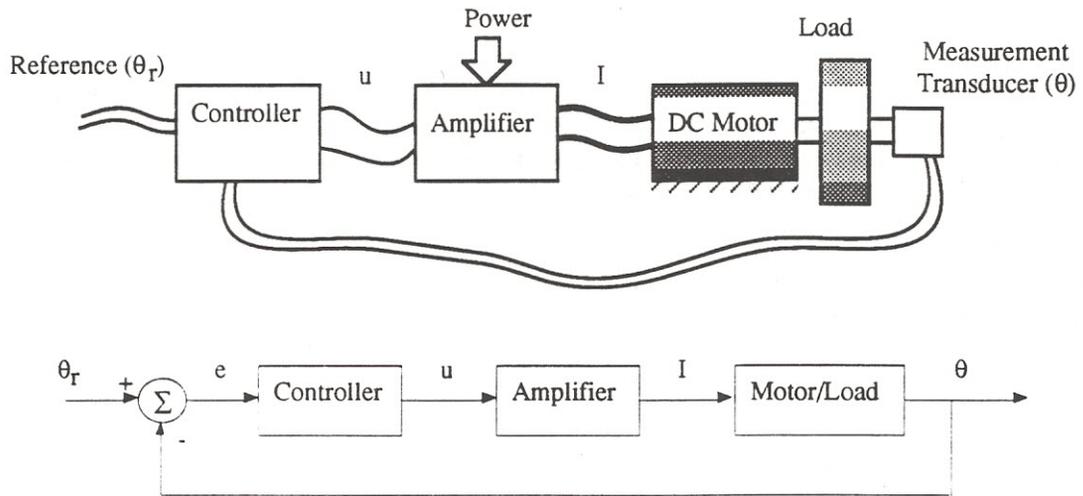


Fig. 2.2 A Simple Position Servo and Equivalent Block Diagram

To begin the analysis of this system, it is first necessary to describe both the static and dynamic behavior of the basic elements, in this case the amplifier, motor and measurement transducer. To do this, we must first develop a basic system dynamics methodology.

## 2.4 Linear Dynamic Processes

Classic control theory is based upon the concept of linear stationary dynamic systems. Such systems are those that are adequately described by ordinary, linear, time-differential equations, with constant (i.e. stationary) coefficients. As an example, consider a simple spring mass system, shown in Fig. 2.3. For this system, the governing equation is derived from a force balance, with appropriate constitutive relations for stiffness and damping:

$$\begin{aligned}
 \Sigma F &= m \ddot{x} = F_{spring} + F_{damper} + F_{ext} \\
 F_{spring} &= -k x \\
 F_{damper} &= -b \dot{x} \\
 m \ddot{x} + b \dot{x} + kx &= F_{ext}
 \end{aligned}
 \tag{2.1}$$

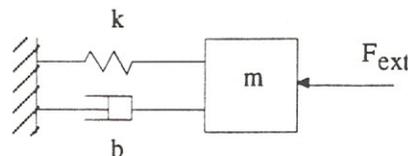


Figure 2.3 Simple Spring Mass Dynamic Systems

As Eqn 2.1 implies, the resulting differential equation is ordinary (no partial derivatives), and linear (all variables and their derivatives are multiplied only by constants). For this physical system we can also illustrate the linearity by looking at the basic constitutive equations for the spring and damper (see Eqn 2.1). Notice that for each one, the force is a linear function of the “input” (either displacement or velocity). These linear elements combine in a linear equation based on Newton’s laws to yield a **linear system**.

Linear differential equations are attractive because they have simple, well behaved solutions, and these solutions have the property of superposition. This latter point is most important to control system analysis, where many and varied types of inputs are often present. With superposition these inputs can be decomposed into sets of standard inputs, for which well known solutions (i.e. outputs) are known. These outputs can then be summed to yield the total system response. The true virtue of linear systems, therefore, is that we can analyze complex systems by artful combinations of various simple systems, and predict the effect of changing these inputs or even the system itself based on this knowledge.

### *Step 1: Modeling the Plant*

The first step is to create a model of the DC motor and load in the servo system of Fig. 2.2. While the motor is a rather complex electromechanical device, the basic constitutive relationship (for a motor with a permanent magnet field and a commutated armature) can be approximated by:

$$T_m = K_t I \quad (2.2)$$

where  $T_m$  is the motor torque,  $I$  is the current in the motor, and  $K_t$  is the motor "torque constant". The torque constant is determined by motor characteristics such as radius, length, number of windings, and field magnet strength. If the motor is operated in a current driven mode (i.e the amplifier in Fig. 2.2 provides a current output in proportion to the actuation input  $u$ ), then it is clear that the motor is a linear device.

What if our concern is not with the torque output of the motor, but with the resulting shaft speed  $\Omega$ ? To determine  $\Omega$ , we must consider the load on the motor, in this case a simple rotational inertia  $J$ , and a bearing viscous damping  $b$ . From this we obtain the dynamics of the mechanical part of the motor system:

$$\Sigma T = J\dot{\Omega} = T_m - b\Omega \quad (2.3)$$

Combining Eqn 2.2 and 2.3 gives the model for the motor with current  $I$  as the input and speed  $\Omega$  as the output:

$$\begin{aligned} J\dot{\Omega} + b\Omega &= K_t I \\ \text{or} \\ (J/b)\dot{\Omega} + \Omega &= (K_t/b)I \end{aligned} \quad (2.4)$$

This is a simple first order linear differential equation.<sup>1</sup> The solution of this equation is well know for several basic inputs. The most common of these is the "step" input, where the current is changed instantaneously from one value to another. The solution to Eqn 2.4 when  $I$  is changed in a stepwise fashion is given by:

$$\Omega(t) = \Omega_{ss} \left[ 1 - \exp\left(-\frac{b}{J}t\right) \right] \quad (2.5)$$

This response is shown in Fig. 2.4, and it can be seen that acceleration (slope) continuously decreases until the steady state velocity, given by the following expression, is reached:

---

<sup>1</sup>Notice that we cannot separate the motor and load because they are coupled dynamically as indicated in this expression. This is a very simple case of preserving input - output *causality* which will become an important issue in separating machine and process in later chapters.

$$\Omega_{ss} = \frac{K_t}{b} I \quad (2.6)$$

Figure 2.4 also illustrates the concept of a *time constant* or characteristic response time. For the exponential response it is defined as the time,  $\tau$ , at which the argument of the exponential is unity. Thus, from Eqn. 2.5, we can define the time constant of the motor as:

$$\tau_m = \frac{J}{b}$$

and Eqn 2.5 can be rewritten for a general output  $y$ :

$$y(t) = y_{ss}(1 - e^{-t/\tau}) \quad (2.7)$$

Equation 2.7 is in fact the general expression for any first order differential equation subjected to a step input, and any physical system that is described by a linear first order differential equation will have this *step response*. However, this simple solution alone is insufficient for designing a control system, since we need a means of concatenating the elements in the control block diagram, and of analyzing the resulting linear system.

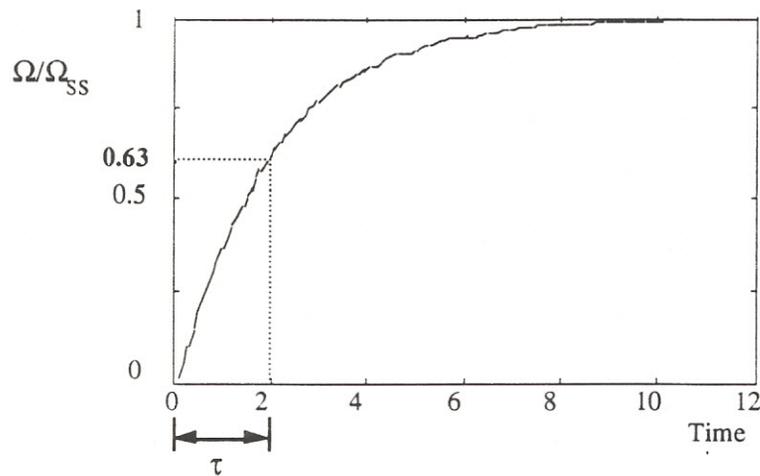


Fig. 2.4 Step Response for a First Order System with a Time Constant = 2.0 sec.

### Describing equations

The block diagram in Fig. 2.1 was constructed both to show the interconnection of each of the elements of the system, and to show the transformation occurring within each block. The contents of each block can be thought of as an **operator** that transforms the input into a corresponding output. For the motor/load block, this transformation is given by Eqn 2.4. However, we need to use a more tractable form than the time differential equation. For this we will transform the differential equations to algebraic equations using operator notation.

The actual mathematics of this transformation are governed by the **LaPlace Transform**<sup>2</sup>. However, for this treatment, we will consider only the use of algebraic operators to illustrate this transformation. For the sake of

<sup>2</sup> The LaPlace Transform of a time function  $y(t)$  is a continuous transform given by the expression:

brevity, it will just be given that the following derivative and integral expressions in  $t$  can be replaced by the corresponding algebraic expressions in  $s$ .

$$\begin{aligned} \frac{d}{dt} \{f(t)\} &\Rightarrow sF(s) - f(0) \\ \frac{d^2}{dt^2} \{f(t)\} &\Rightarrow s^2 F(s) - sf(0) - \dot{f}(0) \\ \int_0^t f(t) dt &\Rightarrow \frac{1}{s} F(s) \end{aligned} \quad (2.8)$$

Any linear, constant coefficient time-differential equation can be transformed into an algebraic equation in  $s$ . Since we can analyze control systems from a “zero” initial state, the initial conditions and integration constants in Eqn. 2.8 vanish, and the most important result is that  $\frac{d}{dt}(f(t)) \Rightarrow sF(s)$ .

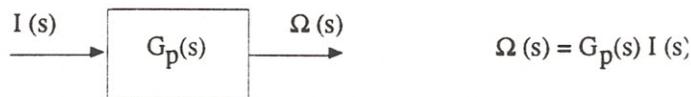
Applying these operators to the motor equation in 2.3 yields:

$$\begin{aligned} J\dot{\Omega}(t) + b\Omega(t) &= K_t I(t) \Rightarrow Js\Omega(s) + b\Omega(s) = K_t I(s) \\ (Js + b)\Omega(s) &= K_t I(s) \end{aligned} \quad (2.9)$$

Notice that the the variables  $\Omega$  and  $I$  are shown as functions of  $s$  rather than time  $t$ , since the transformation has been applied. It is the second expression in Eqn 2.9 that is so powerful, since with a simple manipulation, we can derive a **transfer function** that is in fact the input/output relationship for the motor block:

$$\frac{\Omega(s)}{I(s)} = \frac{K_t}{Js + b} = G_p(s) \quad (2.10)$$

Defining the right hand expression in Eqn. 2.10 as the transfer function  $G_p(s)$ , the motor/load block in Fig 2.2 can be represented in a simple form:



Now we have a compact and algebraically simple means of describing the content of each of the blocks. The content of each block is the transfer function for the process involved, and the output of each block is found simply by multiplying the input by the transfer function.

$$\mathcal{L}\{y(t)\} = \int_{0^+}^{\infty} y(t) e^{-st} dt = Y(s)$$

This transform can be used to solve complex differential equation and a variant of this, the Fourier Transform, is quite useful in the “frequency analysis” of linear system. Further information on the LaPlace Transform in the context of classical control analysis can be found in:

Ogata, Modern Control Principles, Wiley, 1983  
Dorf, Modern Control Systems, Prentice Hall, 1982.

Once the transfer function is defined, it can be analyzed to determine the static and dynamic characteristics of the block. The transfer function for the motor, which is the same as for any **first order system**, can be written in a generic form by using the time constant and steady state gain definitions given above. Thus, a general transfer function for a first order system is:

$$G(s) = \frac{K}{\tau s + 1} \quad (2.12)$$

For the motor example,  $K = K_t/b$  and  $\tau = J/b$ .

The denominator of the transfer function ( $\tau s + 1$ , in this case) corresponds to the left-hand side of the original time differential equation. When the denominator is set equal to zero, it is called the **characteristic equation**, the roots of which correspond to the *inverse* of the characteristic response times of the system. For our simple first order system in Eqn 2.12, the characteristic equation is:

$$\tau s + 1 = 0 \quad (2.13)$$

and the root of this equation is  $-1/\tau$ . It should be apparent from Fig. 2.4 that this one root is sufficient to completely determine the *shape* of the response curve, while the actual magnitude is determined by the size of the input and the steady state gain  $K$ .

Leaving the motor example for a moment, it is evident that higher order differential equations will be encountered in more complex systems. The resulting transfer functions will simply have higher order characteristic equations. As will be shown below, however, linear systems of any order can be decomposed into products of simple first and second order terms. This result follows directly from the observation that any polynomial has (at most) only real and/or complex conjugate roots: the real roots corresponding to first order terms and each complex conjugate pair corresponding to a single second order term.

### ***Control Example: Velocity Servo***

Before attacking the complete position servo of Fig. 2.2, consider controlling the speed of the DC motor. Inspection of the motor transfer function makes this look like a trivial problem, since we need only provide the current  $I$  that will provide the desired speed. This so-called ***open-loop control*** can be applied when the transfer function is completely and perfectly known, but this is seldom the case for any real system. In the case of the motor, "reality" will be in the form of unexpected loads imposed by the environment into which the motor/load system is placed and uncertainty as to the motor constant  $K_t$ , the bearing damping  $b$ , and the load inertia magnitude  $J$ . In this analysis, the unexpected load is modelled as a **disturbance torque** that may arise from friction or other forms of external resistance to motion. It could just as well represent additional sources of power that are undesirable (such as a wind force on a rotating tower, or, if this were a vehicle, the gravitational force from moving downhill).

The updated motor model now has the following form, with the torque disturbance represented by  $T_d$  and time derivatives eliminated by using the operators in Eqn. 2.8.

$$\Sigma T = J\dot{\Omega} = T_m - b\Omega + T_d$$

$$J\dot{\Omega} + b\Omega = K_t I + T_d$$

$$\begin{aligned}
 \Omega(s) &= \frac{K_t}{Js+b} I(s) + \frac{1}{Js+b} T_d(s) \\
 &= \frac{\frac{K_t}{b}}{\frac{J}{b}s+1} I(s) + \frac{\frac{1}{b}}{\frac{J}{b}s+1} T_d(s) \\
 &= G_p(s) \cdot I(s) + \frac{1}{K_t} G_p(s) \cdot T_d(s)
 \end{aligned}
 \tag{2.13}$$

The block diagram for the motor must now be updated to show this new input:

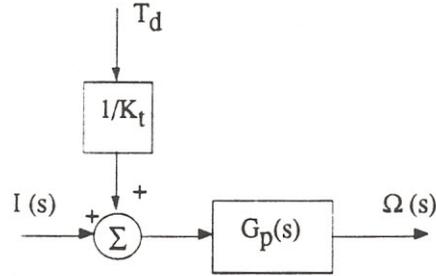


Figure 2.5 Block Diagram of DC Motor with Torque Disturbance

Now imagine the task of controlling the speed  $\Omega$  to be constant when  $T_d$  is unknown and highly variable. Although the motor model in Eqn 2.10 indicates that speed should be proportional to the current (which we could control precisely with a good current amplifier), Eqn 2.13 tells us that the speed will vary directly with the disturbance torque, even if  $I$  is held constant. Clearly we must have some form of *feedback* of the speed so that  $I$  can be manipulated to counteract the effects of  $T_d$ . Borrowing from the canonical feedback system form shown in Fig. 2.1, a velocity control system can be constructed as shown in Fig. 2.6

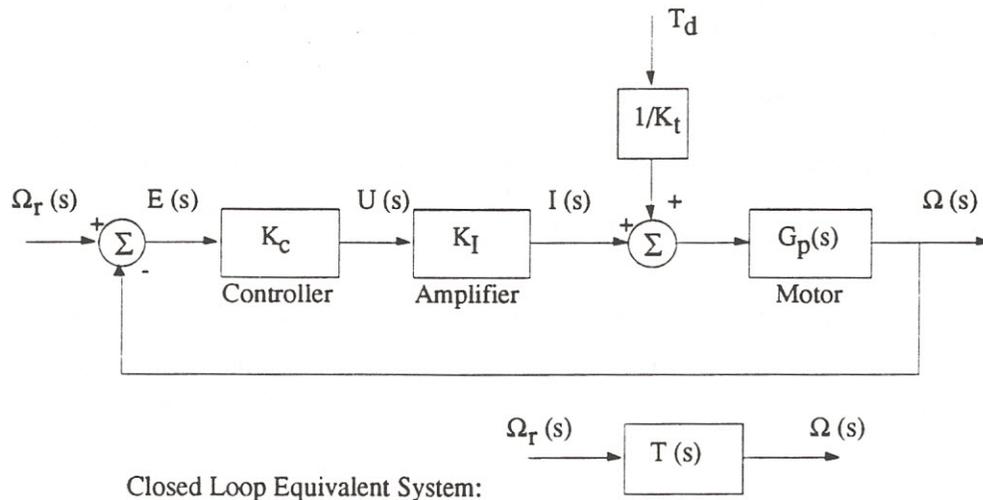


Fig. 2.6 Velocity Servo Block Diagram and Equivalent Closed Loop System

The lower diagram in Fig. 2.6 is a net interpretation of the whole system, with  $T(s)$  as the closed-loop transfer function between  $\Omega$  and  $\Omega_r$ . Note that the motor speed is measured and fed back to a summing junction where it is compared with the desired speed. The resulting error  $E(s)$  is then amplified by the controller gain  $K_c$ , and used to drive a current amplifier, which in turn provides the motor current  $I$ .<sup>3</sup> For simplicity, the amplifier gain  $K_I$  will be assumed to be unity.

Temporarily ignoring the disturbance torque, the closed-loop transfer function for the velocity servo can be found by simple algebraic manipulation based on the block diagram:

$$\begin{aligned}\Omega(s) &= G_p(s) \cdot I(s) \\ I(s) &= K_c(\Omega_r(s) - \Omega(s)) \\ \Omega(s) &= G_p(s) \cdot K_c(\Omega_r(s) - \Omega(s))\end{aligned}\tag{2.14}$$

$$\frac{\Omega(s)}{\Omega_r(s)} = \frac{G_p K_c}{1 + G_p K_c}$$

From the last expression, it is clear that the closed-loop transfer function  $T(s)$  is defined as:

$$T(s) \equiv \frac{G_p K_c}{1 + G_p K_c}\tag{2.15}$$

since it relates the overall system output to the reference or command input.

Substituting for  $G_p(s)$  from Eqn 2.12 yields (and using substitutions  $K_m = K_t/b$  and  $\tau_m = J/b$ ):

$$\begin{aligned}G_p(s) &= \frac{K_m}{\tau_m s + 1} \\ T(s) &= \frac{\frac{K_c K_m}{\tau_m s + 1}}{1 + \frac{K_c K_m}{\tau_m s + 1}} = \frac{\left( \frac{K_c K_m}{1 + K_c K_m} \right)}{\left( \frac{\tau_m}{1 + K_c K_m} \right) s + 1} = \frac{K_T}{\tau_T s + 1}\end{aligned}\tag{2.16}$$

Equation 2.16 still represents a first order transfer function, but with new definitions for  $K$  and  $\tau$ .

A close look at these new system parameters will reveal the properties of this feedback system. Notice first of all that the steady state gain  $K_T$  will never be greater than one, and in fact will only reach unity when  $K_c \rightarrow \infty$ . Thus, this system will never achieve exact command following, but can only approach it as the gains are increased. This apparent drawback is immediately counteracted by the observation that the closed-loop steady-state gain,  $K_T$ , is determined by both  $K_c$  (which we can easily adjust) and  $K_m$  (which is a fixed property of the motor). Thus, if  $K_m$  varies (which it can over time as the components age, or as the temperature of the motor changes) the effect on the steady-state performance of the control system will be slight, provided  $K_c$  is large

<sup>3</sup> While it can be of significance in detailed controller design, we will herein assume the amplifier gain  $K_I = 1.0$ , and that it has no power limitations. Such limits do in fact provide the ultimate limit of performance for many systems, but for now we can assume a "perfect" amplifier.

relative to  $K_m$ . This illustrates the reduction in sensitivity to parameter changes in the system that is achieved by feedback, and this will come to be a most important property as we investigate manufacturing processes where the process gains will be highly variable, and uncertain. Furthermore, we will see below that it is possible to design a controller that can *completely* eliminate the effects of parameter changes at steady-state and provide zero error.

Now consider the effect of feedback on the dynamics of the new system. The closed loop time constant  $\tau_T$  in Eqn 2.16 demonstrates that as the product  $K_c K_m$  is increased, the time constant of the closed loop system will *decrease* leading to a shorter time required to settle after an input change. Thus the effect is clear: the addition of feedback can change the *apparent* time constant of the motor. Theoretically the gain  $K_c$  can be made arbitrarily high to achieve any desired time constant, but the ultimate limit on how fast the motor can respond is typically the power capability of the current amplifier or the heat dissipation capacity of the motor. With this simple example we can see that feedback aids in reducing *variations* in the motor speed and improves the *response time* of the basic motor.

## 2.5 Disturbance Rejection

Recall that the reason for using feedback for speed control was the existence of the disturbance. We can return to the block diagram of Fig. 2.6 and determine the closed-loop *disturbance transfer function* in a similar fashion to Eqn 2.16, but this time ignoring the input  $\Omega_r$ . Performing the block diagram algebra, we obtain:

$$\begin{aligned}\Omega &= G_p \frac{1}{K_t} T_d - G_p K_c \Omega & G_p &= \frac{K_m}{\tau_m s + 1} \\ \frac{\Omega}{T_d} &= \frac{K_m / K_t}{\tau_m s + 1 + K_c K_m} \\ \frac{\Omega}{T_d} &= \frac{\left( \frac{K_m / K_t}{1 + K_c K_m} \right)}{\left( \frac{\tau_m}{1 + K_c K_m} \right) s + 1}\end{aligned}\tag{2.17}$$

Notice that the disturbance transfer function has the same characteristic equation as the velocity controller transfer function in Eqn 2.16 (as it must, since no change has been made to the structure of the *closed-loop* system), but the steady-state gain is considerably smaller. In fact, if  $K_c \rightarrow \infty$ , at steady-state the ratio  $\Omega / T_d \rightarrow 0$ , and the disturbance will be completely rejected. At lower values of gain, the disturbance will be attenuated by the value of  $(K_m / K_t) / (1 + K_c K_m)$ .

As for disturbance dynamics, since the characteristic equation is the same as before, we can expect simple exponential response to disturbances. Specifically, if a step change in  $T_d$  occurs, then a response similar to that of Fig. 2.4 can be expected, but  $\Omega_{ss} = T_d \frac{K_m / K_t}{1 + K_c K_m}$ .

## 2.6 Control for Zero Error

In many control system designs it is desirable to have the error be identically zero. In fact, most manufacturing processes are such that no error introduced by the control system can be tolerated. While it is apparent that high

levels of controller gain  $K_C$  can achieve very low errors, it may not be possible to set the control gain high enough to reach an acceptable error level before certain electrical or mechanical limits are reached.

However, the controller block need not be a simple gain, and we can in fact design this block to be any *realizable* transfer function that is necessary to meet a design goal. In fact, most of today's control systems are implemented using digital computers and the computer can be used directly to implement the transfer function via a set of *difference* equations that approximate the desired dynamic and static behavior of the controller.

For the case of achieving zero error, consider again the velocity servo. The only means of insuring zero error is to let  $K_C$  be infinite, but for zero *steady-state* error there is no requirement that the gain be infinite during transients of the process. Thus, we need a term that gives infinite gain only at steady state. Such performance can be had from a continuous time integral:

$$\int_0^t (\cdot) dt = \frac{1}{s} (\cdot) \quad (2.18)$$

The infinite steady-state gain arises since a constant integrand (e.g. the error signal) will cause this integral to approach infinity as  $t \rightarrow \infty$  (which is in fact the definition of "steady state").

This *integral controller* for the velocity servo is shown in Fig. 2.7

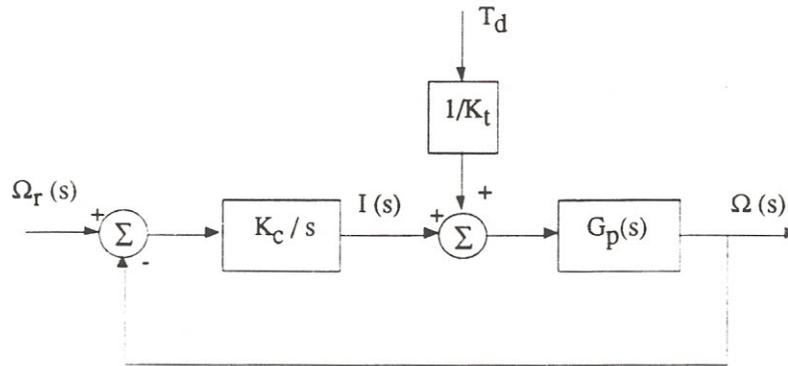


Figure 2.7 Velocity Servo with Integral Control

To determine the net effect of this change on the closed-loop system, we again perform block diagram algebra on this system (ignoring  $T_d$ ), substituting  $K_C/s$  for  $K_C$  in the previous development of Eqn 2.16:

$$T(s) \equiv \frac{G_p \frac{K_C}{s}}{1 + G_p \frac{K_C}{s}}$$

and substituting for  $G_p(s)$ :

$$T(s) = \frac{\frac{K_c K_m}{\tau_m}}{s^2 + \frac{1}{\tau_m} s + \frac{K_c K_m}{\tau_m}} \quad (2.19)$$

Notice now that the denominator is of second order, since the integral controller adds a time dependent element to the system.

To examine the steady state we let all terms in  $s$  vanish (which is equivalent to letting all time derivatives vanish). Now the steady state gain of  $T(s)$  in Eqn 2.19 is unity, *regardless of the values for  $K_c$ ,  $K_m$  or  $\tau_m$* . Thus  $\Omega(s) = \Omega_r(s)$  and the objective of zero steady-state error is achieved.

It is important to examine the effect of this controller on the system dynamics. Clearly there has been a change since the characteristic equation is now of the form ( $b$  and  $c$  are arbitrary constants, no relation to damping  $b$ ):

$$s^2 + \frac{1}{\tau_m} s + \frac{K_c K_m}{\tau_m} = 0$$

or

$$s^2 + b s + c = 0 \tag{2.20}$$

The new dynamics are found by considering the roots of this characteristic equation. Since it is a quadratic, the roots are given by:

$$s_{1,2} = -\frac{b}{2} \pm \frac{\sqrt{b^2 - 4c}}{2} \tag{2.21}$$

and these roots will be either real and negative (if the radical is positive) or complex conjugates (if the radical is negative). If the roots are real, the result is basically two first order roots, and the resulting dynamics will simply be the sum of the response of each of these terms and will resemble the plot shown in Fig. 2.8, which shows the step response of a second order system where the two roots are nearly identical.

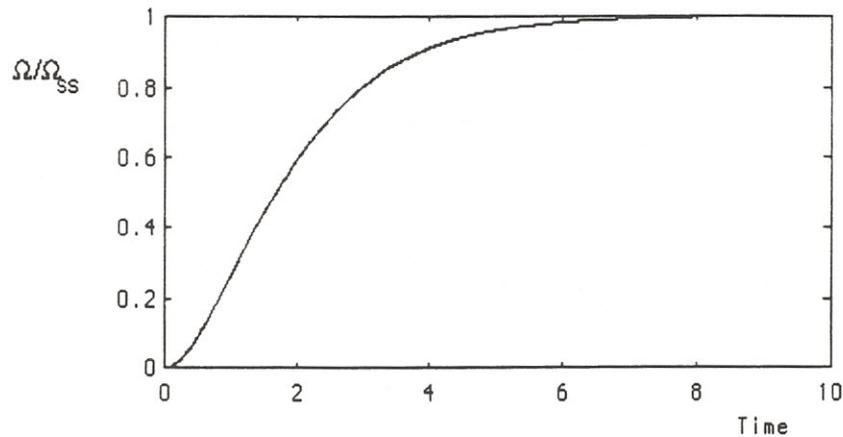


Figure 2.8 Step Response of a Integral Velocity Controller with two Real Roots of Similar Magnitude

However, if the value of the parameters is such that the radical in Eqn 2.21 is negative, the roots will be complex. To better understand the resulting system response, we first make the following substitutions:

$$\begin{aligned}
 s^2 + b s + c &= 0 \\
 \Downarrow \quad \quad \Downarrow \\
 s^2 + 2\zeta\omega_n s + \omega_n^2 &= 0
 \end{aligned}
 \tag{2.22}$$

where the equation has been parameterized by the terms  $\zeta \equiv$  *damping ratio* and  $\omega_n \equiv$  *natural frequency*. The solution of this characteristic equation leads to the step response given by:

$$\Omega(t) = \Omega_{SS} (1 - B e^{-\zeta\omega_n t}) \sin(\omega_d t + \phi)$$

where

$$B = \frac{1}{\sqrt{1-\zeta^2}} \quad \phi = \tan^{-1}\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right) \quad \omega_d = \omega_n \sqrt{1-\zeta^2}
 \tag{2.23}$$

The response will be a sinusoid of frequency  $\omega_d$  (the damped natural frequency), with a magnitude envelope given by the exponential term  $B \exp(-\zeta\omega_n t)$ . This response is illustrated in Figure 2.9 for various values of  $\zeta$ . It is apparent that the natural frequency determines the time scale of the response while the damping ratio determines the shape of the envelope within which the response occurs. (Notice that the rate of decay (or time constant) of this envelope is given by the product  $\zeta\omega_n$ .) Most obvious is the fact that as  $\zeta$  decreases, the degree of *overshoot* beyond the eventual steady-state value, and the degree of oscillation both increase. When  $\zeta = 1$  (critical damping) the response is again that of two simple exponentials, each with equal roots.

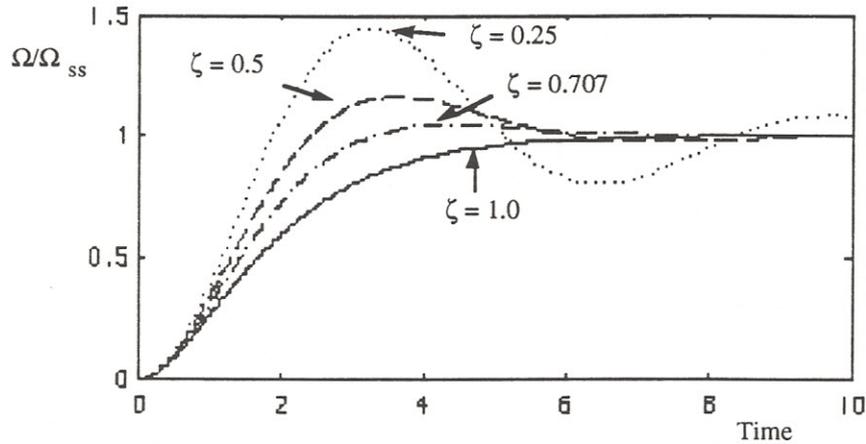


Fig. 2.9 Second Order System Response for Complex Roots

## 2.7 Position Control

The above example dealt only with the problem of *speed control*, whereas we eventually want to control the *position* of a manufacturing machine. Implementation of the position control is quite simple, requiring only that a position measurement be made, as shown in the original servo (Fig. 2.2). The details of this measurement and its actual location in the mechanism are quite important, but for this discussion the ideal case shown in Fig. 2.2 is assumed. For the moment we will also abandon the velocity loop designed above, and concentrate on the position loop only.

The block diagram for the position control system, showing the various transfer functions, is shown in Fig. 2.10. The motor block still has  $\Omega$  as the output, but the position measurement is represented with an integrator block. Thus the measurement of position from a velocity output device ( the motor) adds a dynamic element to the system, and increases the system order by one.

The closed-loop transfer function of the system in Fig. 2.10 is again found by simple block diagram algebra, and is given by:

$$\theta = \frac{1}{s} \frac{K_m}{\tau_m s + 1} u$$

$$u = K_c (\theta_r - \theta) \tag{2.24}$$

$$\frac{\theta}{\theta_r} = \frac{\frac{K_c K_m}{\tau_m}}{s^2 + \frac{1}{\tau_m} s + \frac{K_c K_m}{\tau_m}} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

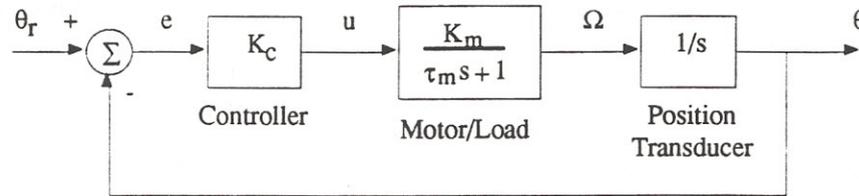


Fig. 2.10 Position Loop Block Diagram : Proportional Control

Note the similarity between Eqn. 2.24 and the closed-loop transfer function for the velocity servo with integral control (Eqn 2.19). In fact they are identical in form, and the only difference is that we added an integrator in the form of the measurement device rather than in the controller. (It will become evident later, that *where* the integrator is added is quite important, but for now the distinction is lost.)

Since this system has the same dynamics (i.e. the same denominator) as the integral controller velocity servo, the step response plot will be of the same character as those of Fig. 2.9, but now the output is, of course,  $\theta$  rather than  $\Omega$ .

Now consider the role of the controller gain  $K_c$ . from Eqn 2.24. Using the variable substitutions given earlier we can assign:

$$\omega_n^2 = K_c K_m / \tau_m \tag{2.25}$$

and

$$2\zeta\omega_n = 1/\tau_m \tag{2.26}$$

From the last expression we can see that the **damping**, which is defined as  $2\zeta\omega_n$ , is a constant. This means that regardless of how  $K_c$  is varied, the envelope of the response will remain constant. Since the envelope determines the **settling time** of the transient,  $K_c$  will have no effect on how rapidly the system achieves a new equilibrium, this factor being completely determined by the motor time constant  $\tau_m$ .

So what effect does  $K_c$  have? The answer is found by examining the expression for  $\omega_n^2$  (Eqn 2.25), where it is apparent that  $\omega_n \propto \sqrt{K_c}$ . Also, re-examining Eqn 2.26 shows that:

$$\zeta = \frac{1}{2\omega_n\tau_m} = \frac{1}{2\sqrt{K_c}K_m\tau_m} \quad (2.27)$$

or that  $\zeta \propto 1/\sqrt{K_c}$ . So the effect of *increasing*  $K_c$  is to *increase*  $\omega_n$  and *decrease*  $\zeta$ , leading to a variety of responses (shown in Fig. 2.11 for various  $K_c$ , with  $K_m$  and  $\tau_m$  assumed to be unity) that increase in overshoot and oscillation frequency as  $K_c$  is increased, but which all settle to steady state at essentially the same time.

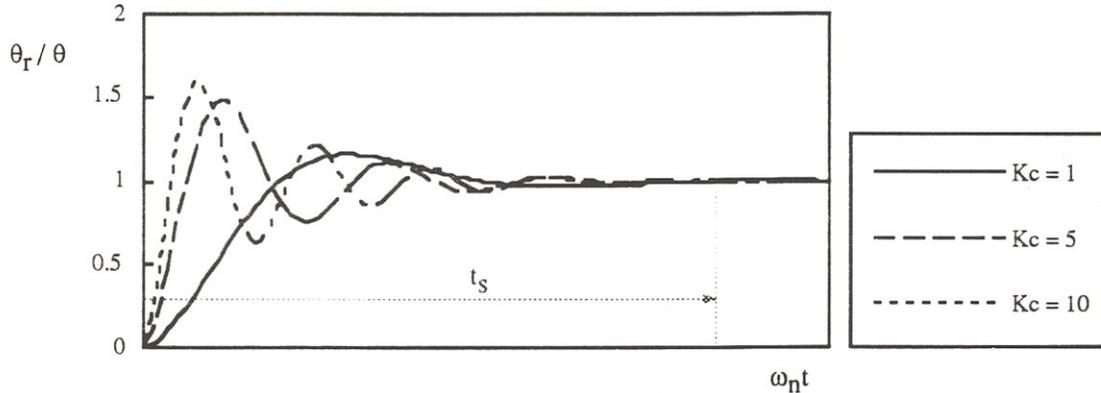


Fig. 2.11 Step Responses for the Proportional Position Servo  
Note the calling on the abscissa and the definition of the settling time  $t_s$ .

Note also from Fig. 2.11 that the the settling time is essentially the same for all three responses and that both the time and amplitude axes have been non-dimensionalized. For the time axis the normalizing factor is  $\omega_n$ , and it is thus apparent that the two parameters for a second order system :  $\zeta$  and  $\omega_n$ , determine the degree of overshoot and the time scaling of the response respectively.

#### *Damping Control in the Position Servo: Velocity Feedback*

The response of Fig. 2.11 is not very satisfactory, especially if the major design criterion is settling time. The implication is that if a shorter settling time is desired, a new motor/load system must be specified, which negates one major benefit of control, that is altering the closed-loop system without the need for altering the process itself. To overcome this problem, it is necessary to gain access to the first order term of the characteristic equation for this system, since that term defines the system damping.<sup>4</sup>

<sup>4</sup> Notice that if we had such control, we could determine *all* of the coefficients in the characteristic equations, and thus could specify completely the system dynamics. In the theory of **state-determined systems** this is equivalent to having full state control, that is we can modify all the coefficients of the characteristic equations using feedback control. Such **state feedback systems** are typical of what is referred to as Modern Control Systems and represent a different approach to feedback control than discussed in this Chapter. For More information on State Feedback see:

Ogata, Modern Control System Principles, Chapter 16  
Kwakernak and Sivan, State Variable Control Systems

While there are several ways of motivating our approach, consider a simple physical argument. Damping is defined as a force acting in opposition to the motion and (if the system is linear) acting in proportion to the velocity. Thus damping is proportional to the negative of the velocity. Examining Fig. 2.10, it is evident that if in addition to measuring and feeding back position, we as well measure and feedback (with a negative sign) the velocity  $\dot{\theta}$  then we should add damping to the system. Thus we have modified the system as shown in Fig. 2.12.

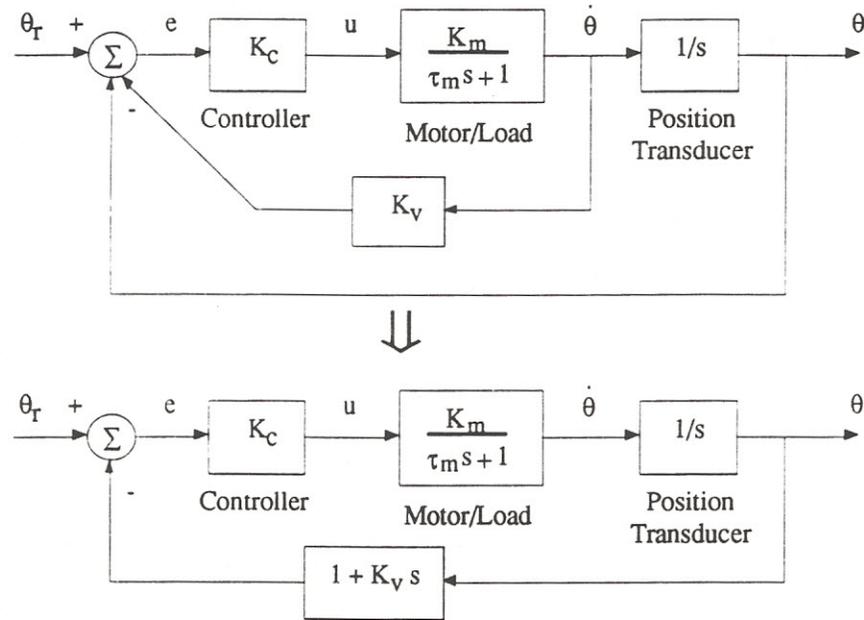


Figure 2.12 Position Servo with Velocity Feedback and a Block Diagram Equivalent

From this figure we can rederive the closed loop transfer function:

$$\theta = \frac{K_c K_m}{s(\tau_m s + 1)} e$$

$$e = \theta_r - (1 + K_v s)\theta$$
(2.28)

$$\frac{\theta}{\theta_r} = \frac{\frac{K_c K_m}{\tau_m}}{s^2 + \frac{(1 + K_v K_c K_m)}{\tau_m} s + \frac{K_c K_m}{\tau_m}}$$

Again making the variable substitutions we can solve for:

$$\omega_n^2 = \frac{K_c K_m}{\tau} \quad (\text{unchanged from before})$$

and

(2.29)

$$\zeta = \frac{1 + K_v K_c K_m}{2 \sqrt{K_c K_m \tau_m}} \quad (\text{new parameter } K_v)$$

The effect of the velocity feedback is now obvious;  $\omega_n$  can be arbitrarily specified by  $K_C$  and  $\zeta$  can then be adjusted to any desired value with  $K_V$ .

### Limits to Performance : Unmodeled effects

This design implies that *any* speed of response is possible, and indeed that is true, but only up to the power capabilities of the amplifier and motor combination, as mentioned earlier. To find the ultimate limit of this system, it is necessary to reexamine the assumptions made earlier. These include the simplified model of the motor, the modeling of the amplifier as “perfect” and the implication that the measurements were perfect and “noise-free”. In fact, both of these effects: model errors and measurement noise will be dominant limits when addressing any real system, and in particular when applying this theory to manufacturing processes.

The effect of model errors is quite complex, and would require a more in-depth treatment of control theory. However, it is clear that if the process is modeled incompletely, then the design derived by the transfer function analysis given above will not produce the expected performance. However, it may produce acceptable performance *over a specified range of operation*. Defining this range is an important aspect of control system design, and some rudiments of this specification are covered below under the heading of frequency response. Most importantly, modeling errors can lead to stability problems, where the performance is not only different from expectations, but also unstable and potentially disastrous. Therefore, knowledge of *stability analysis* is also essential to control system design.

Measurement error has a direct and easily quantified effect on the system. The controller can only regulate the system to the perceived output value, thus any measurement errors will be directly reflected into output errors under feedback control. The effect of noise is more subtle, and requires a closer definition of noise. First, we can characterize noise as any unwanted signal that does not relate to the actual physical variable being sensed. However, if this noise is a *zero mean process*, then it is often acceptable to assume that on average the noise signal is zero. In such a case the noise does not affect the mean value of the output, but only causes variations about that mean. If it is not zero mean, then it again produces a steady-state measurement error that cannot be averaged out by the control system.

### Disturbance Rejection for the Position Servo

Finally we must consider the basic problem of *regulation*, where the control objective is to maintain the output at the reference value in the face of unwanted disturbances. Borrowing from the velocity servo analysis, we can model the disturbance as shown in Fig. 2.5, and then simply augment the block diagram of Fig. 2.12, as shown in Fig. 2.13.

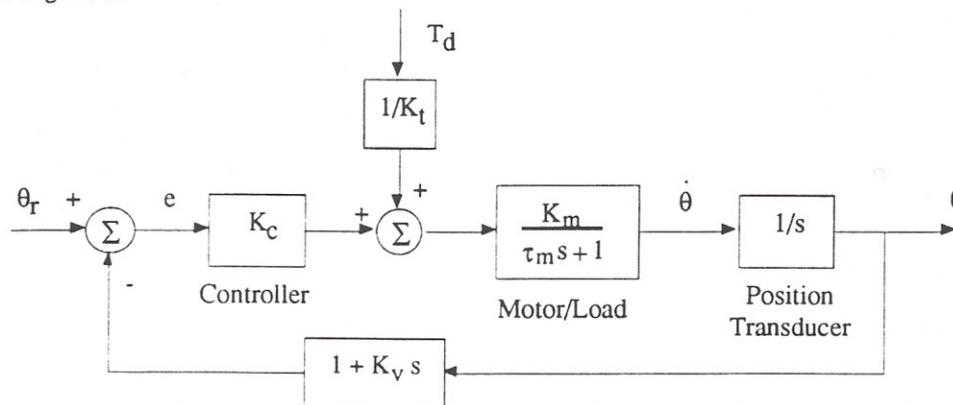


Figure 2.13 Position Servo with Torque Disturbance

We again set the reference to zero and concentrate on deriving a disturbance transfer function, which not surprisingly (given Eqn 2.17) is found to be:

$$\theta = \frac{K_m}{s(\tau_m s + 1)} (-K_c(1 + K_v s) \cdot \theta + T_d / K_t) \quad (2.30)$$

$$\frac{\theta}{T_d} = \frac{(K_m / K_t) / \tau_m}{s^2 + \frac{(1 + K_v K_c K_m)}{\tau_m} s + \frac{K_c K_m}{\tau_m}}$$

As before we now find that the disturbance is reduced by a steady-state factor  $1/K_c K_m$ , and that as the control gain  $K_c$  is increased, the effect of the disturbance is diminished. But why has it not been eliminated, since there is a “free integrator” in the loop, just as with the example in Fig. 2.7 for the velocity servo?

### Proportional - Integral (PI) Control

This result indicates that the location of the disturbance and integrator in the loop is quite important, and suggests that to eliminate  $T_d$  completely, we must again incorporate an Integrator into the controller. However, this time we will do so by *adding* the integrator to the controller rather than exchanging it for the proportional term. Thus, the controller block will be  $G_c = K_c + K_I/s$  or a **P**roportional + **I**ntegral (PI) controller. PI controllers are quite common and are used precisely for the purpose of providing zero steady-state error to a system that is otherwise “OK”.

With this new controller substituted into the block diagram of Fig. 2.13, the disturbance transfer function becomes:

$$\begin{aligned} \frac{\theta}{T_d} &= \frac{(K_m / K_t) / \tau_m}{s^2 + \frac{(1 + K_v K_c K_m)}{\tau_m} s + \frac{K_m (K_c + K_I / s)}{\tau_m}} \\ &= \frac{s(K_m / K_t) / \tau_m}{s^3 + \frac{(1 + K_v K_c K_m)}{\tau_m} s^2 + \frac{K_m K_c}{\tau_m} s + \frac{K_m K_I}{\tau_m}} \end{aligned} \quad (2.31)$$

Since the numerator of Eqn 2.31 vanishes as  $s \rightarrow 0$ , the PI controller has eliminated the steady-state effect of the disturbance. Since steady-state implies that  $s \rightarrow 0$ , then in steady-state  $\theta/T_d = 0$ . Note also that all three coefficients in the characteristic equation can be adjusted independently using  $K_v$ ,  $K_c$ , and  $K_I$ .

In addition to the error elimination, the familiar second order characteristic equation is gone, and we are faced with a third order denominator. How can we again decompose this equation into the familiar  $\tau$ ’s,  $\omega_n$ ’s and  $\zeta$ ’s? In fact we simply determine the roots of the denominator and produce the factored characteristic equation:

$$(s + p_1)(s + p_2)(s + p_3) = 0 \quad (2.32)$$

where the roots  $p_i$  can be real or complex. We know that at least one must be real (since complex roots must have conjugates), so there are two possibilities: all three roots are real:

$$(s + p_1)(s + p_2)(s + p_3) = 0 \quad (2.33)$$

or one real root and one pair of complex roots:

$$(s + p_1)(s^2 + 2\zeta\omega_n s + \omega_n^2) = 0 \quad (2.34)$$

Detailed analysis of these roots is beyond this treatment, but we can speculate on expected responses. Linear system theory, and the Laplace Transform can be used to show that the response of these higher order systems will in fact be the sum of responses of each factor. The relative weight applied to each factor is a function of their relative speed and the nature of the numerator (which itself can have dynamic terms). Figures 2.14 and 2.15 illustrate this effect for Eqns 2.33 and 2.34. Note in 2.15 that when one of the roots is much larger than the others, its effect is negligible.

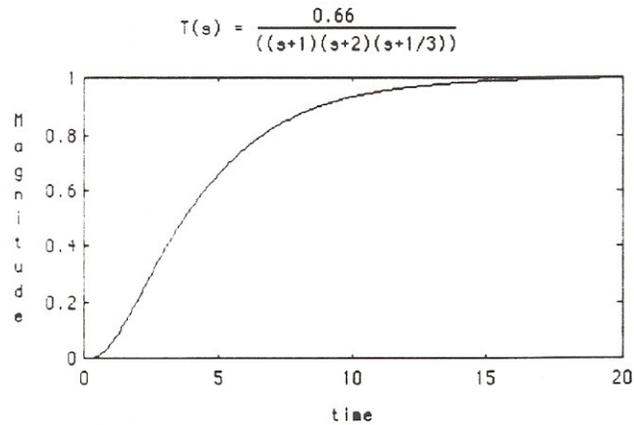


Fig. 2.14 Third Order System Response with Real Roots

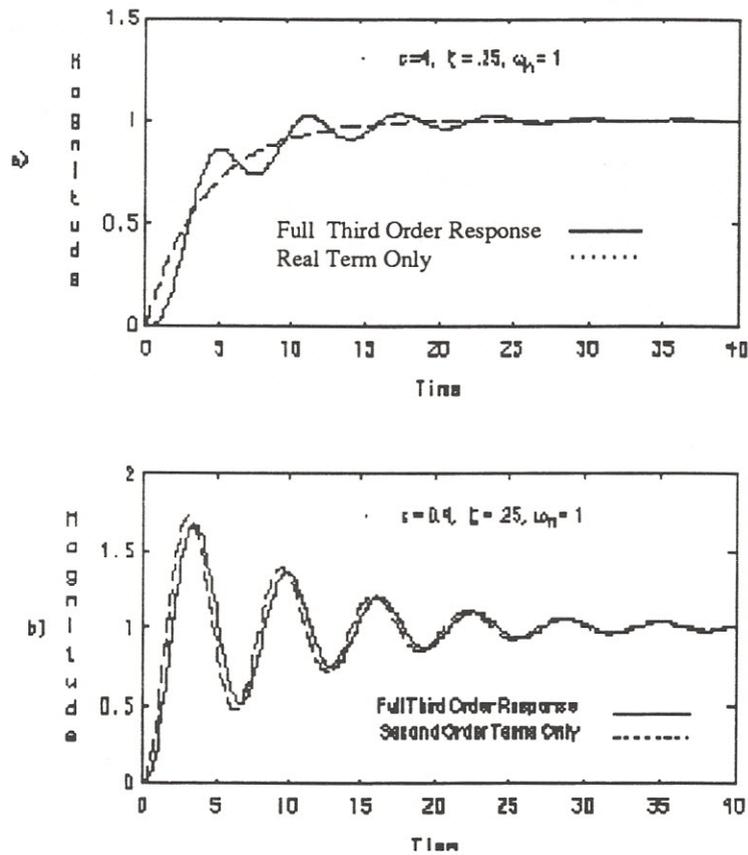


Fig. 2.15 Third Order System Responses for the Closed Loop Transfer Function

$$T(s) = \frac{\omega_n^2 / \tau}{(s + 1/\tau)(s^2 + 2\zeta\omega_n s + \omega_n^2)}$$

- a) "slow" real root
- b) "fast" real root

## 2.8 Speed, Accuracy, Bandwidth: System Response

The analysis so far has concentrated on the time-response of the servo. From the step response plots, a measure of response speed (the settling time) has emerged. This is an intuitively appealing measure, but it does not fully describe how the system will respond to a less standard signal, in particular a time-varying one, applied either at the input or through the disturbance. The disturbance especially may present a periodic signal whose frequencies could be numerous and over a broad band. As a result, it is useful to consider a *frequency domain* view of the control system, and to ascertain the *frequency response* of the system. It is also quite convenient to use this description to design the best control system for selective disturbance rejection and noise elimination.

Frequency response is simply the *steady - state* response of a system to a sinusoidal input. One very useful property of linear systems is that when subjected to sinusoidal excitation, the output is itself a sinusoid, always of the same frequency. However, the amplitude and phase of the output relative to the input is typically quite

different, as shown in Fig. 2.16, and it is this difference that provides a wealth of information about the system.<sup>5</sup>

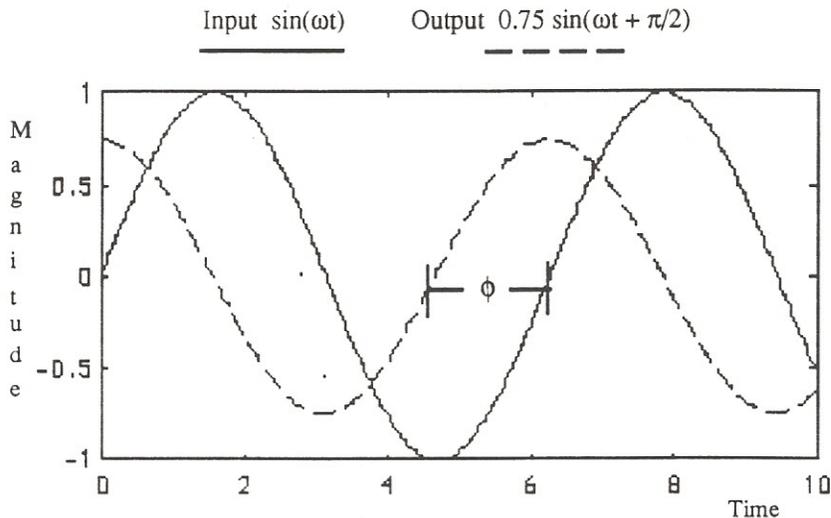


Fig. 2.16 Input Sine Wave and Linear System Output

Notice that the output differs only in the magnitude and phase angle, but not the frequency

To analyze a system described by transfer functions, we make the substitution  $s = j\omega$  and then follow the rules of complex number algebra. Fortunately this need only be done for a few simple elements, and the rest of the analysis follows from combination of these simple elements.

Consider again a simple first order system (e.g. the DC motor). The *sinusoidal transfer function* for such a system becomes:

$$G(s)|_{s=j\omega} = G(j\omega) = \frac{K}{\tau \cdot j\omega + 1} \quad (2.35)$$

In this  $j\omega$  domain, the input is implied to be a sinusoid of frequency  $\omega$  and unity amplitude (although it can be scaled linearly to any value). Since  $G(j\omega)$  is a complex number, it can be described in terms of orthogonal components (real and imaginary parts) or in polar form, with a magnitude and phase angle. For  $G(j\omega)$  these two precisely describe the amplitude and phase of the response to a sinusoid. Thus for Eqn 2.35:

<sup>5</sup> Proof of this analytically can be done by using the LaPlace Transform. Consider the example of a first order system  $G(s) = \frac{K}{s+a}$ . Taking the transform of the sine input:  $\mathcal{L}(\sin(\omega t)) = \frac{\omega}{s^2 + \omega^2} = R(s)$ . The output is then found as  $C(s) = G(s)R(s)$ . Thus

$$C(s) = \frac{K\omega}{(s+a)(s^2 + \omega^2)}$$

Using the *inverse transform*, the time response  $c(t)$  is given by:

$$c(t) = \alpha e^{-j\omega t} + \bar{\alpha} e^{+j\omega t} + b e^{-at}$$

Since our analysis is only for steady state, the first term vanishes, leaving us with a complex expression:

$$c(t) = \alpha e^{-j\omega t} + \bar{\alpha} e^{+j\omega t} = \beta \sin(\omega t + \phi)$$

where  $\alpha$  and  $\beta$  can be shown (see Ogata chapter 9) to be functions of the parameters of  $G(s)$  and the frequency  $\omega$ .

$$|G(j\omega)| = \frac{K}{\sqrt{\tau^2\omega^2 + 1}} \quad (2.36)$$

$$\angle G(j\omega) = \tan^{-1} \frac{\text{Im}(G)}{\text{Re}(G)} = -\tan^{-1}(\tau\omega) \quad (2.37)$$

A quick look at Eqn 2.36 indicates that as the frequency of the input is *increased*, the output magnitude will *decrease*, and Eqn 2.37 indicates that the phase angle will be negative (a phase “lag”) and will become more negative with frequency.

Since both  $|G|$  and  $\angle G$  are functions of  $\omega$ , they can be plotted on similar scales to illustrate the effect of *all* values of frequency on the system. In addition, as will be shown later, it is best to plot both variables on a log  $\omega$  scale and to plot log magnitude as well. The result is a **Bode Diagram** and Fig. 2.17 shows a non-dimensional Bode Diagram for a simple first order system with  $K = 1$  and  $\tau = 1$ .

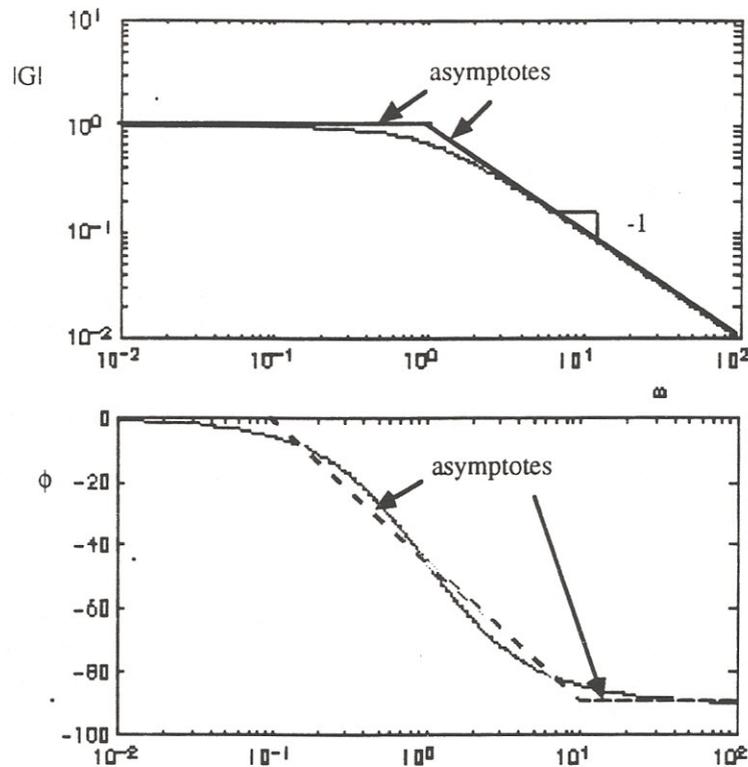


Fig. 2.17 Bode Diagram for  $G(s) = \frac{1}{s+1}$

We can interpret this plot by looking at how the magnitude and phase change with  $\omega$ . Notice that at very low frequencies  $|G|$  (often referred to as the “gain”) is unity and remains nearly at that level until near  $\omega = 1/\tau$ . Beyond  $\omega = 1/\tau$ , the gain drops off rapidly, eventually along a straight line of slope -1 on the log-log scale. Notice also that at low  $\omega$  the phase is nearly zero, but as the point  $1/\tau$  is approached, the *lag* approaches  $-45^\circ$  and eventually reaches an asymptote at  $-90^\circ$ . The significance of the point  $1/\tau$  is more evident when considering asymptotic approximations to this plot. As shown on Fig. 2.17, both the magnitude and phase plots are well approximated by straight-lines. For the gain these lines are in fact the low frequency and high frequency asymptotes. For the phase, the low and high frequency asymptotes are horizontal lines at  $0$  and  $-90^\circ$ , and the

phase change is well approximated by a line from  $0^\circ$  at  $\omega = 0.1/\tau$  to  $-90^\circ$  at  $\omega = 10/\tau$ . Thus, knowing only the time constant of the transfer function, we can immediately plot the Bode Diagram using asymptotic approximations. Notice that any gain term ( $K$ ) in the transfer function is applied as a simple vertical scaling of the magnitude plot, but has no effect on the phase.

Now consider the second order underdamped system:

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

Rewriting slightly, and making the substitution  $s = j\omega$ , the sinusoidal transfer function becomes:

$$G(j\omega) = \frac{1}{(1 - \frac{\omega^2}{\omega_n^2}) + 2j\zeta \frac{\omega}{\omega_n}} \quad (2.38)$$

Again looking at magnitude and phase of this complex ratio:

$$|G(j\omega)| = \frac{1}{\sqrt{\left[1 - \left(\frac{\omega}{\omega_n}\right)^2\right]^2 + \left[2\zeta \frac{\omega}{\omega_n}\right]^2}}$$

$$\angle G(j\omega) = -\tan^{-1} \frac{2\zeta \frac{\omega}{\omega_n}}{1 - \left(\frac{\omega}{\omega_n}\right)^2} \quad (2.39)$$

A Bode Diagram for this transfer function is shown in Fig. 2.18. As can be seen from the transfer function, as  $\omega \Rightarrow 0$  again  $|G| \Rightarrow 1$ , and as  $\omega \Rightarrow \infty$ ,  $|G| \Rightarrow 1/\omega^2$ , or a slope of -2 on the Bode Diagram. Again the major changes occur at  $\omega = \omega_n$ , and the asymptotic gain plot shows that the change in slope from 0 to -2 occurs at that so-called **break point**. While the asymptotic plot is quite simple and appears identical to the first order plot, save for the increased slope, a closer look at the diagram shows the effect of the damping ratio. As the frequency approaches  $\omega_n$ , the magnitude *increases* and reaches a peak at a point near  $\omega_n$  (the point of resonance is  $\omega_r = \omega_n \sqrt{1 - 2\zeta^2}$ ). While the asymptotic plot misses this effect, on either side of the resonance the agreement is quite good.

The phase also looks somewhat like a simple doubling of the first order case, and the asymptotic plot is simple a line from  $0^\circ$  at  $\omega = 0.1\omega_n$  to  $-180^\circ$  at  $\omega = 10\omega_n$ . However, notice that the actual phase change becomes more severe as the damping ratio decreases.

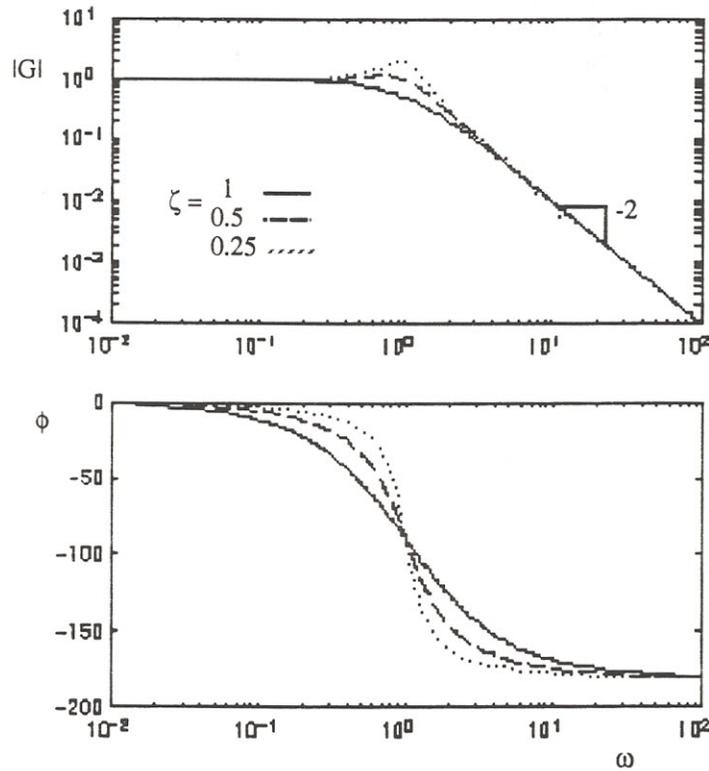


Fig. 2.18 Bode Diagram for  $G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$

### Basic System Characterization : Bandwidth

How do plots of the frequency response help in design? There are many specific methods associated with the design of feedback control systems based on use of the Bode Diagram, but here we want only to introduce a few major concepts that will be useful when dealing with manufacturing processes.

The first concept is that of *bandwidth*. Simply defined, the bandwidth is the range of frequencies over which the system response is “useful”. There can be many definitions of “useful”, but the most common is that the output is useful as long as it is greater than 0.5 to 0.707 times the input. While this may seem arbitrary, and an integer (such as one!) may seem more convenient, it is chosen because of the asymptotic character of most systems.

Consider the simple first order system. The gain diagram shown in Fig. 2.17 shows that if we chose  $|G| \geq 1$  as the definition, the bandwidth would also be zero, since the magnitude never actually reaches unity. However, if we choose 0.707, this point in fact occurs precisely at the break point where  $\omega = 1/\tau$ . Thus the bandwidth of this simple first order system is from 0 to  $1/\tau$ . Likewise, for our unity gain second order system shown in Fig. 2.18, the bandwidth is well approximated by the range 0 to  $\omega_n$ . Notice that the bandwidth will in fact be variable depending upon  $\zeta$ , and will be precisely equal to  $\omega_n$  only when  $\zeta = 1.0$ .

To consider the real utility of the bandwidth measure, it is necessary to again look at a feedback system. For this we again turn to the servo example. The question arises: if we know the frequency response of the motor alone, what is the frequency response of the proportional speed servo of Fig. 2.6 and how does  $K_C$  affect this response? The answer requires some complex algebra, which leads to a reduced solution.

General block diagram algebra on a closed-loop system, such as that shown in Fig. 2.19, reveals that the closed-loop transfer function is given by:

$$T(s) = \frac{G(s)}{1+G(s)H(s)}$$

and if  $H(s) = 1$

$$T(s) = \frac{G(s)}{1+G(s)} \quad (2.40)$$

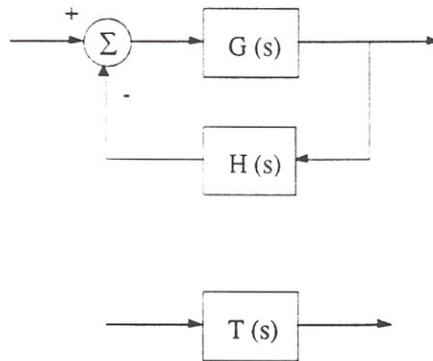


Fig. 2.19 A General Block Diagram and Equivalent Closed-Loop System

To find the gain plot for the closed loop velocity servo, we need only substitute the appropriate transfer function, set  $s=j\omega$ , and then determine the magnitude. For the general closed-loop transfer function with unity  $H(s)$ :

$$\begin{aligned} T(j\omega) &= \frac{G(j\omega)}{1+G(j\omega)} \\ |T(j\omega)| &= \left| \frac{G(j\omega)}{1+G(j\omega)} \right| \\ &\approx \frac{|G(j\omega)|}{1+|G(j\omega)|} \end{aligned} \quad (2.41)$$

The last expression, which is an approximation, makes the problem quite simple, since we already know the term  $|G|$  from the bode diagram. From Eqn 2.41 we can see that if  $|G| \gg 1$ , then  $|T|$  tends to 1, and as  $|G|$  approaches 0,  $|T|$  tends to zero. Finally, if  $|G| = 1$ , then  $|T| \approx 0.5$ , which happens to be one measure of useful output from our bandwidth definition. In other words, if the open-loop transfer function gain  $|G|$  is much greater than one, the closed-loop gain  $|T|$  will always approach unity.

Now to our servo problem. The  $G(s)$  (forward loop transfer function) for our servo is given by:

$$G(s) = \frac{K_c K_m}{\tau_m s + 1} \quad (2.42)$$

The Bode Diagram for this transfer function is simply that of Fig. 2.17 scaled by the gains  $K_c K_m$ . The resulting gain diagram is shown in Fig. 2.20 along with the corresponding plot for  $|T|$ . To determine the corresponding closed-loop bandwidth, we really only need to find the point at which  $|G|$  is unity. This is defined as the **crossover point** (or crossover frequency). Fig. 2.20 shows the crossover point for  $K_c K_m = 10$ , but in principle  $K_c$  can be adjusted to any desired value. If we define the bandwidth of this closed-loop system, as the range over which the output is  $\geq 0.5$ , then it is clear that the crossover point ( $\omega_c$ ) on the  $|G(\omega)|$  plot corresponds to the closed-loop bandwidth  $\omega_b$ .

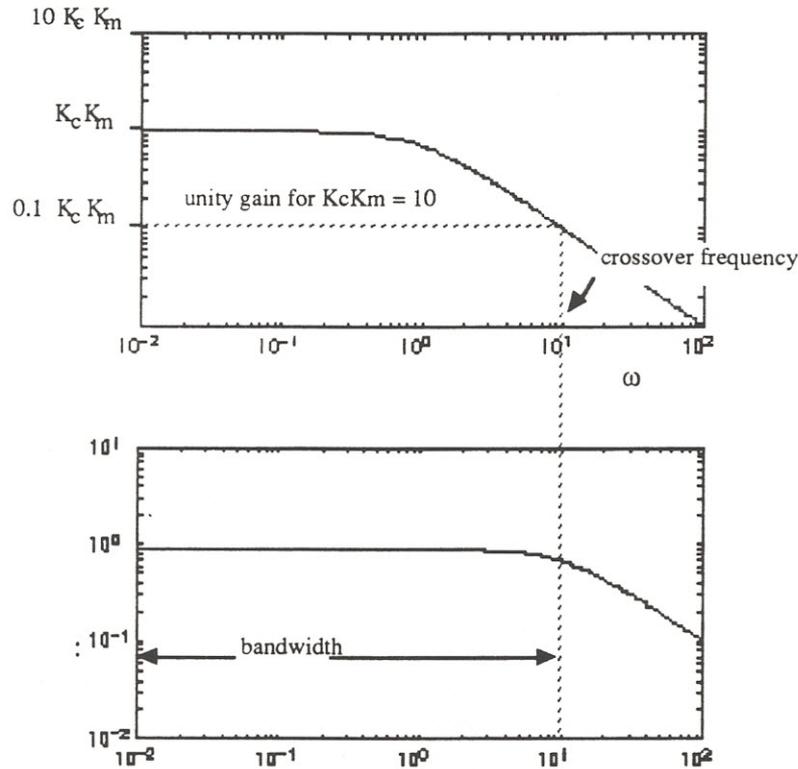


Fig. 2.20 Open-Loop / Closed-Loop Correspondence for a Velocity Servo

Thus the bode diagram for the open-loop system can be used directly, without further plotting or analysis, to determine the crossover frequency, which in turn tells us the closed-loop bandwidth.

### *Effect of Measurement Noise and Disturbances*

Returning to the simple velocity servo problem, consider the existence of noise in the feedback path, as shown in Fig. 2.21, arising from the measurement process. This noise enters the loop in a manner similar to the input, so the noise transfer function is identical to the input-output transfer function. Therefore, the frequency response to noise in the measurement is the same as the closed-loop diagram in Fig. 2.20. Now assume that the noise is characterized by the signal  $n(t) = \sin(\omega t)$ . From Fig. 2.20, it can be seen that if the noise frequency is less than  $\omega_b$ , the noise amplitude will be reflected to the output with no attenuation. However, if the noise frequency is greater than  $\omega_b$ , it will be attenuated by the factor  $1/(\omega - \omega_b)$ . This leads to the well known conclusion that noise above the bandwidth of the system is usually not a problem (until we try to increase the bandwidth!) but noise within the bandwidth of the system is a major problem. The Bode Plot allows quick, graphical evaluation of noise rejection in this manner.

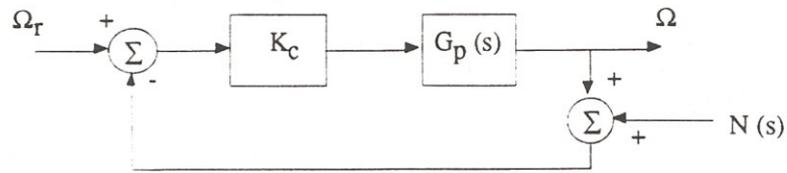


Fig. 2.21 Measurement Noise in a Velocity Servo

If a disturbance enters the system, as shown in Fig 2.22, the

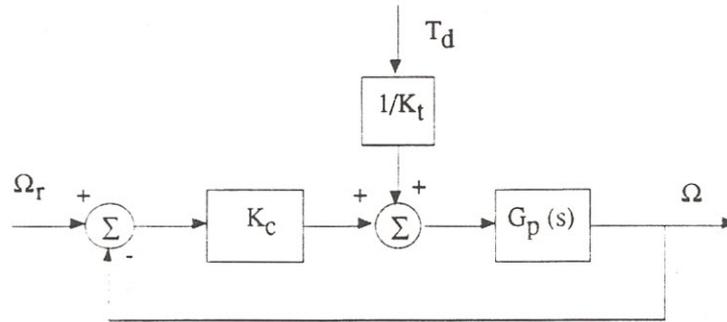


Fig. 2.22 Velocity Servo with a Disturbance

disturbance transfer function can be found again from block diagram algebra as ( recall  $G_p = K_m/(\tau_m s + 1)$  ):

$$\frac{\Omega}{T_d} = \frac{G_p/K_t}{1 + K_c G_p} \quad (2.43)$$

The Bode Diagram for this transfer function could be drawn knowing only  $G_p$  and  $K_c$ , but we can in fact simply notice the difference between this expression and the closed loop transfer function

$$\frac{\Omega}{\Omega_r} = \frac{K_c G_p}{1 + K_c G_p}$$

$$\frac{\Omega}{T_d} = \left( \frac{1}{K_c K_t} \right) \frac{\Omega}{\Omega_r} \quad (2.44)$$

Thus the disturbance frequency response is simply the closed loop response scaled by the factor  $(1/K_c K_t)$ . Near zero frequency the rejection is  $(K_m/K_t)/(1 + K_c K_m)$ , which is precisely what was found earlier for the steady-state disturbance rejection of this servo. Note that for this particular disturbance, which enters the system at the input to the plant, high frequency disturbances will be rejected, simply because the bandwidth of the motor itself (which is  $1/\tau_m$ ) will prevent any high frequency inputs (such as the disturbance) from propagating to the output.

### ***Closed-Loop System Stability: Right-Half Plane Roots and Phase margin***

All of the systems discussed above have had clear and robust points of equilibrium. This can be traced to the analytical fact that the real parts of all the roots of the characteristic equations (which, recall, are the arguments of exponentials in the time response functions) were all negative. But what if through the use of feedback, which we know can alter the roots of the "new" closed-loop characteristic equation, one or more roots become

zero or positive. This implies that at least one exponential term in the time solution will diverge with time, and approach infinity. Such a condition would be unstable, and must be avoided. Not only can we use linear system theory to detect instability, we can also use it to develop a measure of proximity to stability limits, or the margin of stability.

To illustrate the concept of root change and instability, consider the third order system developed earlier when PI control was applied to the position servo (Eqn 2.27). The characteristic equation for that system is:

$$s^3 + \frac{(1+K_v K_c K_m)}{\tau} s^2 + \frac{K_m K_c}{\tau} s + \frac{K_I K_m}{\tau} = 0 \quad (2.45)$$

Assuming for the moment that  $K_c$  and  $K_v$  are fixed, and we will only vary  $K_I$ , this can be simplified to:

$$s^3 + a_2 s^2 + a_1 s + K = 0 \quad (2.46)$$

where  $K = K_I \frac{K_m}{\tau}$ .

By solving for the roots of Eqn 2.46 as  $K$  is varied from 0 to  $\infty$ , and then plotting these on the complex or "s"-plane, the location of the roots for all possible values of gain can be easily seen. This plot, referred to as the root-locus for the closed-loop system, is shown in Fig. 2.23, assuming the values shown for  $K_c$ ,  $K_v$ , and with  $K_m = 1.0$ .

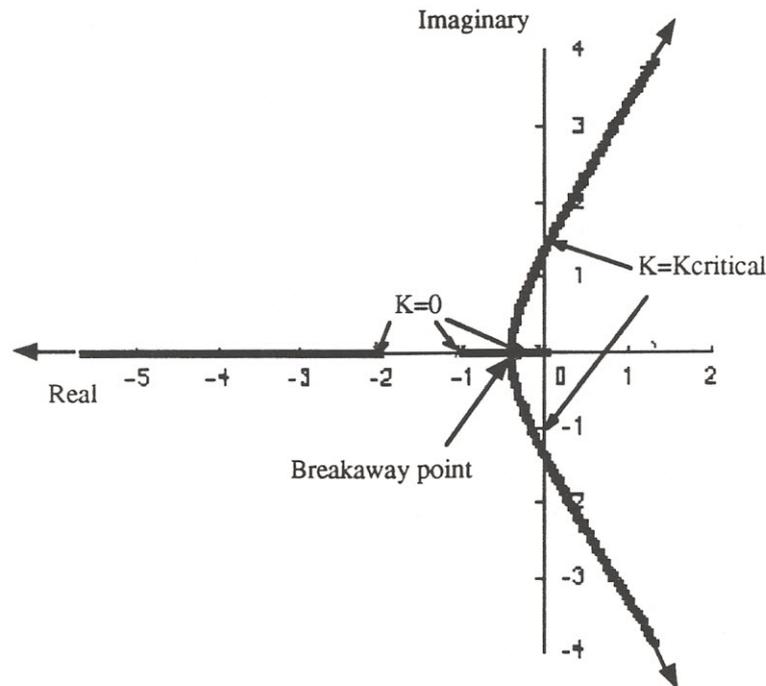


Figure 2.23 Locus of Roots in the s-Plane for the System of Eqn 2.42 with  $K_c = 2$ ,  $K_v = 1$ , and  $\tau = 1$

To interpret this plot, recall that if a root is real, it has simple exponential response. Thus any root that lies on the abscissa has no complex part, and the magnitude represents  $1/\tau$  of a first order term in the response. Any

root that appears in the general s-plane has both real and complex parts, and is therefore an oscillatory root pair. The rate of decay of such terms is determined by the real-part of the root.<sup>6</sup>

With this rapid introduction to s-plane analysis, Fig. 2.23 becomes more clear. At  $K = 0$ , we have essentially an open loop system and the roots are the roots of the open-loop block in the system. Note that the integrator appears as a root at the origin. As  $K$  is increased, these roots move in such a way that one always remains real (as expected) and the other two converge on the real axis and then “break” into the complex plane. Thus as the gain  $K$  is increased, a pair of complex roots develops. As  $K$  is increased further, the real root becomes faster and faster (and therefore less significant to the response) and the complex roots change by moving farther from the origin ( $\omega_n$  increasing) and by moving closer to the imaginary axis (settling time increasing). Finally, there is a **critical gain** at which point the complex roots lie on the imaginary axis. This implies zero damping or sustained oscillation for these roots. This point is called **marginally stable**. As  $K$  is increased, the roots move into the positive real half-plane (or right-half plane) and diverge, leading to an unstable system.

The existence of any gain value that causes right-half plane roots indicates the potential for instability, and the proximity of roots to the imaginary axis is a measure of relative stability. The more negative the real part of the roots the more stable the system. This corresponds to the observation that the more negative the real part the faster the settling time. Thus relative stability and settling time are directly related.

Stability can also be analyzed using the Bode Diagram. Proof of frequency domain stability is beyond the treatment of this text, and can be found in any of the texts listed in the bibliography at the end of this Chapter. However, an intuitive argument can be made to back up the analytical result.

Consider a linear closed-loop system of high order (i.e.  $>2$ ). Further assume that the phase lag of the forward loop transfer function  $G$  is exactly at  $-180^\circ$ , and that  $|G|$  is exactly 1. The net effect of this is to make  $G$  appear like a simple negation block. This in turn cancels the negative sign at the summing junction, and we are left with what is known as a **positive feedback** system. Since errors will cause control actions that reinforce rather than diminish the error, positive feedback systems are **unstable**.

Thus a critical point in the frequency response of the open-loop transfer function is that of **unity gain and  $-180^\circ$  phase**. Proximity to this point is again an indication of a margin of stability.

To illustrate this concept consider the third order system shown in Fig. 2.24 (a simple 3 term real system) The Bode Diagram for this system can be easily constructed using graphical addition. Since the individual first order terms multiply together to form the Open-Loop transfer function  $G(s)$ , they can be plotted individually on the Bode Diagram and then simply **added**. This is true for the magnitude since we use the log and it is true for the phase since multiplication of complex numbers in polar form requires addition of the angle terms.

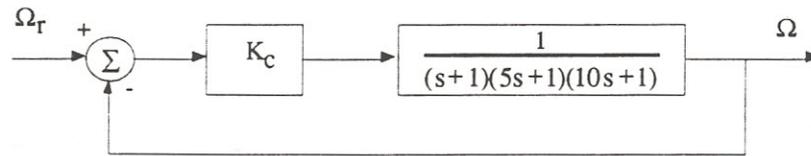


Fig. 2.24 Third Order System for Frequency Study

The individual terms are plotted in Fig. 2.25, and the sum of these is shown with the solid line. Most notable in this plot is the fact that the phase angle now goes from 0 to  $-270^\circ$ . Based on the stability criterion stated above, this means that instability will occur, if the gain  $|G|$  were to be at or greater than unity when the phase

<sup>6</sup> It can easily be shown as well that the distance from the origin to the complex root is  $\omega_n$ . Further, the angle this vector makes with the real-axis is defined as  $\phi$ , then the damping ratio for the complex roots can be shown to be  $\cos^{-1}(\phi)$ .

$< -180^\circ$ . If the gain  $K$  in the system is raised (as indicated by downward motion of the unity gain line), then clearly this condition could occur at higher values of the gain. This is consistent with the root locus result in Fig.2.23, which likewise indicated instability at higher values of control gain.

Thus stability in frequency analysis relates to the proximity to a unity gain  $-180^\circ$  phase condition. If the control gain of the system is adjusted such that this condition is avoided, then a stable, converging response is expected. Further, the margin of stability can be quantified in two ways:

$$\text{Phase Margin} = \phi_m = 180^\circ + \angle G \big|_{|G|=1}$$

$$\text{Gain Margin} = K \big|_{\phi=-180}$$

Thus by simple inspection of the Bode Diagram for any particular choice of control gain, we can ascertain the relative stability of the system. Phase margin is particularly useful since for a second order system it can be shown that  $\zeta \approx 0.01 \phi$ , and this approximation is quite good even for system of higher order.

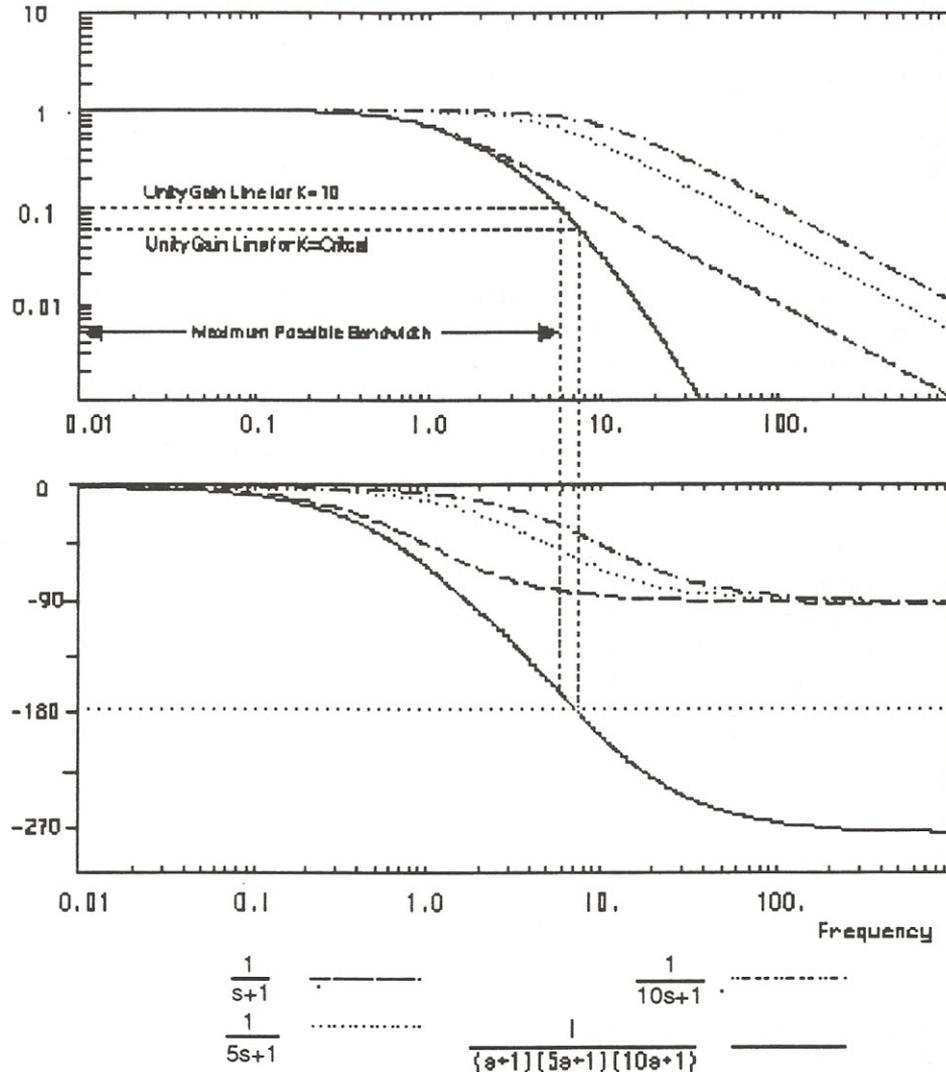


Fig. 2.25 Bode Diagram of the System In Fig. 2.24

### Effect of Measurement Delays

A prevalent problem in any manufacturing process is access to measurable outputs. In many cases a compromise must be made whereby measurements are taken some time after the output actually occurs. Quality Control sampling is an extreme example of this. In a faster time scale, measurements are often delayed by physical phenomena that intervene between the point of production and the point of measurement. In either case, a significant *pure delay* is introduced that can have a profound effect on the performance of the feedback system.

To analyze this effect, it is most convenient to return to the Bode Diagram. A pure time delay term can be expressed in input-output terms as:

$$y(t) = u(t - T_d)$$

where  $T_d$  is the delay time.

In the  $s$ -domain this "block" can be expressed by the transfer function:

$$G_d(s) = e^{-Ts} \quad (2.47)$$

Now consider a simple velocity servo with a pure delay in the feedback loop, as shown in Fig. 2.26. To develop the Bode Diagram for this system, we plot a first order term for the motor, and again represent  $K_C$  as a scale shift in  $|G|$ . To add the delay effect, we must first look at it in terms of gain and phase:

$$G_d(s)|_{s=j\omega} = e^{-j\omega T} = |G|e^{-\angle G}$$

Thus

$$|G_d| = 1$$

$$\angle G = \phi = -\omega T \quad (2.48)$$

This tells us that a pure delay will introduce additional phase lag into the system, and rather severely so, since it is proportional to the frequency, and thus has no asymptotic value.

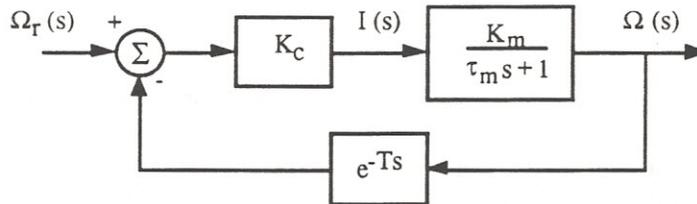


Fig. 2.26 Velocity Servo with a Measurement Delay

The Bode Diagram in Fig. 2.27 shows the effect of a 0.1 sec delay term on the velocity servo. Without the delay we know that no stability limit exist for the velocity servo. Both our earlier analysis and the observation that the phase never goes below  $-90^\circ$  proves this point. However, when the phase lag of the delay is added, it is apparent that above  $\omega = 19$  rad/sec the phase plummets below  $-180^\circ$  and puts a firm limit on the maximum possible gain. From this diagram we can see that the critical gain  $K_{crit} \approx 20$ . It is also apparent from Eqn 2.48 that if  $T_d$  increases, this limit will become even smaller.

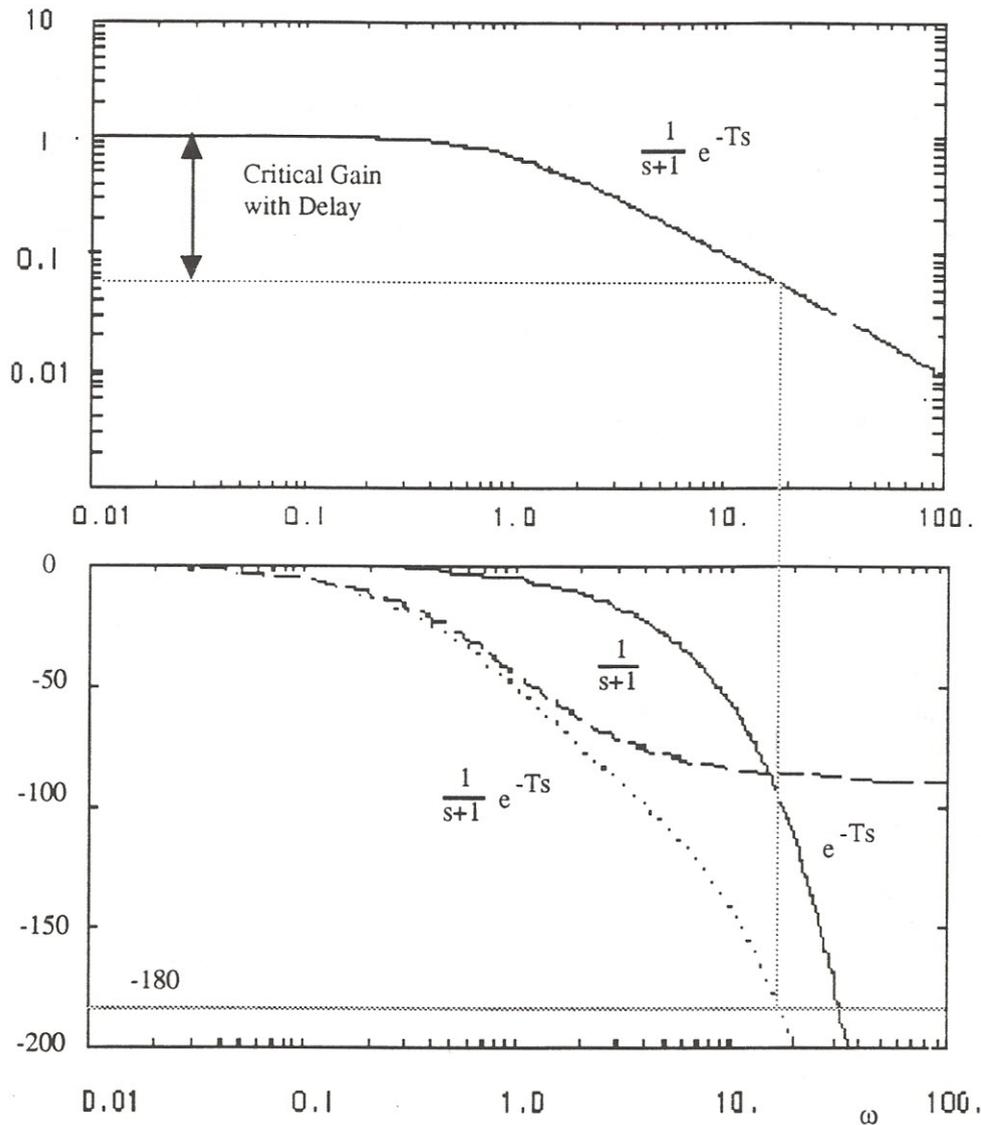


Fig. 2.27 Effect of a Pure Delay on Phase of a Second Order System

When the gain is limited so is the crossover frequency, which in turn limits the closed-loop bandwidth of our servo. Thus the most immediate effect of the delay on any system is to place a limit on the bandwidth that can be achieved. This in turn limits the speed of response. However, note that this also limits the amount of control gain that can be applied. Recalling that this gain directly affects the disturbance rejection and error characteristics of most systems, limiting the gain magnitude has the effect of impairing the disturbance-rejection and error-eliminating properties of the controller. While these problems can be mitigated using free-integrators, this in turn only adds more phase lag to the system <sup>7</sup>.

<sup>7</sup> A free integrator is essentially a first order system with zero time constant, and simple analysis shows that  $|G| = 1/\omega$  and  $\phi = -90^\circ$ . Thus a free integrator adds a constant  $-90^\circ$  of phase lag, thereby decreasing any phase margin and further limiting the potential bandwidth. Notice, however, that  $|G| \rightarrow \infty$  as  $\omega \rightarrow 0$ , just as was shown to be necessary for zero steady-state errors.

## 2.9 Discrete or Digital Control Systems

In virtually all implementations of feedback control, the digital computer is used as both the controller and the error junction, and often as a feedback filter as well. In some cases the application is done in such a way that the substitution of a discrete device (the computer) into the theory for continuous system is inconsequential. However, for most applications this is not the case and there exists a complete theory for discrete systems that parallels most of what has been presented here, with certain “new” results added.

Figure 2.28 shows one type of digital control system, presented again in the context of the position servo.

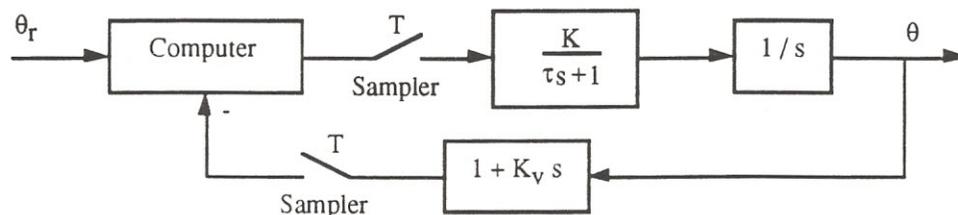


Figure 2.28 A Canonical Digital Control System

Notice that the continuous signals from the feedback measurement are connected to the computer with a *sampler*. This device takes “snapshots” of the signals, and then presents these to the computer at regular time intervals or “samples”  $T$ . Herein lies the major departures from the continuous systems we have dealt with so far. Because of the sampler (which is necessary owing to the serial, stepwise nature of computers) the controller is given fragmentary data that is only updated after a time  $T$ . The simple proportional control algorithm used in the position servo earlier would be implemented in some type of line code in the computer, requiring steps such as:

```

Read Input Data  $D$ 
Error = Reference -  $D$ 
 $U = K * \text{Error}$ 
Write Output  $U$ 

```

In the simplest of implementations, the time required to execute these steps would determine the sample time  $T$ , since input data could only be read that frequently. Note also that there is a pure delay of  $T$  between when the input is read and when the output is sent.

Without further analysis it should be evident that a digital controller will differ from a continuous one at least by the addition of a new parameter, the sample time  $T$ . Indeed it is this sample time that governs much of the differences, and much of the performance limits for digital controllers. If  $T$  is very small (i.e. we are using a very fast computer), the difference between the continuous and discrete systems can be neglected. However as  $T$  gets large, the intervals between updates can be long enough to permit significant changes to occur in the system, and the performance can degrade considerably.

But how do we judge whether  $T$  is large or small? Clearly it relates to the speed of the system we are controlling. Thinking instead of a sample frequency  $1/T$ , the concept of bandwidth is quite helpful. If  $1/T < \omega_b$  significant changes in the system can occur between samples, and it is quite possible for the controller to do exactly the opposite from desired, since its information is so delayed. As  $1/T$  approaches  $\omega_b$  this situation lessens, but we are still faced with uncertainty. Consider the classic problem shown in Fig. 2.29. Here a simple sine wave is being sampled, and presented to a computer for “interpretation”. Note that depending upon the phase at which it is sampled, the data could reflect anything from a zero input to a constant, to a sine wave of lower frequency. This problem is nearly alleviated when  $1/T$  becomes twice the sine wave frequency, and when  $1/T$  is greater than twice the sine wave frequency the problem disappears.

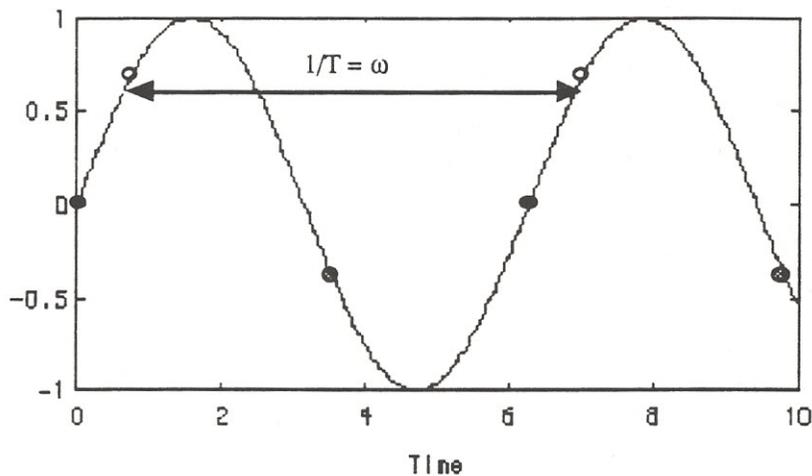


Fig. 2.29 Effect of Sampling on Interpretation of a Sine Wave.  
 Note that sample at  $1/T = \omega$  would indicate that the wave is a constant,  
 with amplitude dependent upon the phase of the samples.

Sample time also has a profound effect on stability. Recall that the simple proportional control code requires a time  $T$  to execute. This delay will therefore add phase lag to the system, even if it is of low order. As shown in Fig. 2.26 this phase lag will have a significant effect on phase margin (and therefore damping and stability) as we either increase the bandwidth of the system or increase the delay time. Thus the sample time has a direct effect on limiting the maximum bandwidth of our control system, and can also lead to decreased relative stability (e.g. damping) even for low order systems.

The theory of discrete systems is well developed, and well beyond the scope of this text. Several excellent texts on this topic are listed in the bibliography of this chapter.

## 2.10 Summary

This Chapter has attempted to cover the material contained in complete texts and cannot, of course, be considered a complete exposure to classical control theory. However, several basic principles have been illustrated, including:

- The definition of a linear system**
- Transfer function analysis of linear systems**
- Feedback control**
- Steps in control system design:**
  - Modeling**
  - Controller design**
  - Response prediction and/or simulation**
- Frequency response**
- Bandwidth**
- Stability of closed-loop systems**

We have used the velocity and position servo as the model example, and have seen the properties of a simplified version of such a system. While our analysis has been necessarily rudimentary, it does reflect the basic character of servos, and more complex models and controller design can be built directly upon this work.