# Chapter 6

# Transfer Functions

In this chapter we introduce the concept of a *transfer function* between an input and an output, and the related concept of block diagrams for feedback systems.

## 6.1 Frequency Domain Description of Systems

The idea of studying systems in the frequency domain is to characterize a linear time-invariant system by its response to sinusoidal signals. The idea goes back to Fourier, who introduced the method to investigate propagation of heat in metals. Frequency response gives an alternative way of viewing dynamics. One advantage is that it is possible to deal with systems of very high order, even infinite. This is essential when discussing sensitivity to process variations. This will be discussed in detail in Chapter **??**.

Frequency response also gives a different way to investigate stability. In Section 2.3 it was shown that a linear system is stable if the characteristic polynomial has all its roots in the left half plane. To investigate stability of a the system we have to derive the characteristic equation of the closed loop system and determine if all its roots are in the left half plane. Even if it easy to determine the roots of the equation numerically it is not easy to determine how the roots are influenced by the properties of the controller. It is for example not easy to see how to modify the controller if the closed loop system is stable. The way stability has been defined it is also a binary property, a system is either stable or unstable. In practice it is highly desirable to have a notion of the degrees of stability. All of these issues can be related to frequency response. The key is Nyquist's stability criterion (the subject of the next chapter) which is a frequency response concept.

Frequency response was one of the key ideas that formed the foundation of control.

The response of linear systems to sinusoids was discussed in Section 2.3, see Equation (2.22). Consider a linear input/output system

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + \ldots + a_n y = b_1 \frac{d^{n-1} u}{dt^{n-1}} + b_2 \frac{d^{n-2} u}{dt^{n-2}} + \ldots + b_n u, \qquad (6.1)$$

whose characteristic equation has eigenvalues $\lambda_k$. The output corresponding to the input $u(t) = \sin \omega t$ can be shown to be of the form

$$y(t) = \sum_k C_k(t) e^{\lambda_k t} + |G(i\omega)| \sin(\omega_t + \arg G(i\omega)),$$

where the coeffiecients $C_k$ depend on the parameters of the differential equation. If the system is stable, i.e. $\mathrm{Re}(\lambda_k) < 0$ for all $k$, the first term will decay exponentially and the solution will converge to the steady state response given by

$$y(t) = |G(i\omega)| \sin(\omega_t + \arg G(i\omega)) \qquad (6.2)$$

where $G(i\omega)$ represents the gain of the system and $G(i\omega)$ its phase. (The reason for multiplying the frequency by $i$ will be made clear later.) This is illustrated in Figure 6.1 which shows the response of a linear time-invariant system to a sinusoidal input. The figure shows the output of the system when it is initially at rest and the steady state output given by (6.2). The figure shows that after a transient the output is indeed a sinusoid with the same frequency as the input.

## 6.2 Transfer Functions

The model (6.1) is characterized by two polynomials

$$\begin{aligned} a(s) &= s^n + a_1 s^{n-1} + a_2 s^{n-2} + \ldots + a_{n-1} s + a_n \\ b(s) &= b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_{n-1} s + b_n \end{aligned}$$

The rational function

$$G(s) = \frac{b(s)}{a(s)} \qquad (6.3)$$

is called the transfer function of the system.

Consider a system described by (2.6) assume that the input and the output have constant values $u_0$ and $y_0$ respectively. It then follows from (2.6) that

$$a_n y_0 = b_n u_0$$

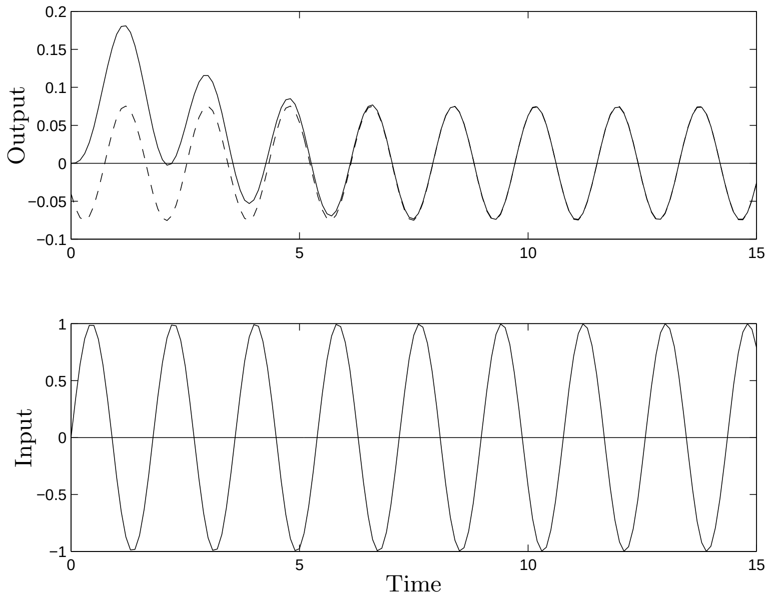


Figure 6.1: Response of a linear time-invariant system to a sinusoidal input (full lines). The dashed line shows the steady state output calculated from (6.2).

which implies that

$$\frac{y_0}{u_0} = \frac{b_n}{a_n} = G(0)$$

The number $G(0)$ is called the static gain of the system because it tells the ratio of the output and the input under steady state condition. If the input is constant $u = u_0$ and the system is stable then the output will reach the steady state value $y_0 = G(0)u_0$. The transfer function can thus be viewed as a generalization of the concept of gain.

Notice the symmetry between $y$ and $u$. The *inverse system* is obtained by reversing the roles of input and output. The transfer function of the system is $\frac{b(s)}{a(s)}$ and the inverse system has the transfer function $\frac{a(s)}{b(s)}$.

The roots of $a(s)$ are called poles of the system. The roots of $b(s)$ are called zeros of the system. The poles of the system are the roots of the characteristic equation. If $\alpha$ is a pole it follows that $a(\alpha) = 0$ and that $y(t) = e^{\alpha t}$ is a solution to the homogeneous equation (2.7). To prove this we differentiate

$$\frac{d^k y}{dt^k} = \alpha^k y(t)$$

and we find

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + a_2 \frac{d^{n-2} y}{dt^{n-2}} + \ldots + a_n y = a(\alpha) y(t) = 0$$

If $s = \alpha$ is a pole the solution to the differential equation has the component $e^{\alpha t}$, which is also called a mode, see (2.15). The modes correspond to the terms of the solution to the homogeneous equation (2.7) and the terms of the impulse response (2.17) and the step response.

If $s = \beta$ is a zero of $b(s)$ and $u(t) = Ce^{\beta t}$, then follows that

$$b_1 \frac{d^{n-1} u}{dt^{n-1}} + b_2 \frac{d^{n-2} u}{dt^{n-2}} \ldots + b_n u = b(\beta) C e^{\beta t} = 0.$$

If the input to the system (2.6) is $e^{\beta t}$ it thus follows that the solution is given by the general solution to the homogeneous equation (2.7). The solution thus will not contain the term $e^{\beta t}$. A zero of $b(s)$ at $s = \beta$ blocks the transmission of the signal $u(t) = Ce^{\beta t}$. Notice that this does not mean that the output is zero when the the input is $e^{\beta t}$ unless initial conditions are chosen in a very special way.

### Transfer Function of a State Space System

Consider a linear state space system of the form

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx. \end{aligned}$$

We know from the previous chapter that the solution of this system can be written using the convolution integral

$$y(t) = Ce^{At} x(0) + \int_0^t Ce^{A(t-\tau)} B u(\tau)\, d\tau.$$

It is easy to show that if the system is stable with $x(0) = 0$ and $u(t)$ is a sinusoid, then $y(t)$ is also a sinusoid. We can thus ask to find the transfer function for this linear system.

To find the input/output relation, we differentiate the output to obtain

$$\begin{aligned} y &= Cx \\ \frac{dy}{dt} &= C\frac{dx}{dt} = CAx + CBu \\ \frac{d^2y}{dt^2} &= CA\frac{dx}{dt} + CB\frac{du}{dt} = CA^2x + CABu + CB\frac{du}{dt} \\ &\vdots \\ \frac{d^ny}{dt^n} &= CA^nx + CA^{n-1}Bu + CA^{n-2}B\frac{du}{dt} + \ldots + CB\frac{d^{n-1}u}{dt^{n-1}} \end{aligned}$$

Let $a_k$ be the coefficients of the characteristic equation. Multiplying the first equation by $a_n$, the second by $a_{n-1}$ etc we find that the input-output relation can be written as

$$\frac{d^ny}{dt^n} + a_1\frac{d^{n-1}y}{dt^{n-1}} + \ldots + a_ny = B_1\frac{d^{n-1}u}{dt^{n-1}} + B_2\frac{d^{n-2}u}{dt^{n-2}} + \ldots + B_nu,$$

where the matrices $B_k$ are given by.

$$\begin{aligned} B_1 &= CB \\ B_2 &= CAB + a_1CB \\ B_3 &= CA^2B + a_1CAB + a_2CB \\ &\vdots \\ B_n &= CA^{n-1}B + a_1CA^{n-1}B + \ldots + a_{n-1}CB \end{aligned}$$

Using a bit more linear algebra, it can be show that the resulting transfer function is simply

$$G(s) = C[sI - A]^{-1}B. \tag{6.4}$$

We illustrate this with an example.

*Example* 17 *(Transfer Function of Inverted Pendulum).* The linearized model of the pendulum in the upright position is characterized by the matrices

$$A = \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}, \quad B = \begin{matrix} 0 \\ 1 \end{matrix}, \quad C = 1 \quad 0, \quad D = 0.$$

The characteristic polynomial of the dynamics matrix $A$ is

$$\det(sI - A) = \det \begin{matrix} s & -1 \\ -1 & s \end{matrix} = s^2 - 1$$

Hence

$$(sI - A)^{-1} = \frac{1}{s^2 - 1} \det \begin{matrix} s & 1 \\ 1 & s \end{matrix}$$

The transfer function is thus

$$G(s) = C[sI - A]^{-1}B = \frac{1}{s^2 - 1} 1 \quad 0 \begin{matrix} s & 1 \\ 1 & s \end{matrix}_{-1} \begin{matrix} 0 \\ 1 \end{matrix} = \frac{1}{s^2 - 1}$$

### Poles and Zeros

The poles of the linear time-invariant system (**??**) are simply the eigenvalues of the matrix $A$. To determine the zeros we use the fact that zeros are such that the pure response to input $e^{st}$ is zero. The state and the output that corresponds to the input $u_0 est$ are then $x_0 e^{st}$ and $y_0 e^{st}$ where

$$\begin{aligned} sx_0 &= Ax_0 + Bu_0 \\ y_0 &= Cx_0 + Du_0 \end{aligned}$$

Requiring that the output is zero we find

$$\begin{matrix} sI - A & Bx_0 \\ C & Du_0 \end{matrix} = 0$$

This equation has a solution with nonzero $x_0$, $u_0$ only if the matrix on the left has nonzero rank. The zeros are thus the values $s$ such that

$$\det \begin{matrix} sI - A & B \\ C & D \end{matrix} = 0 \tag{6.5}$$

Notice in particular that if the matrix $B$ has full rank the matrix has $n$ linearly independent rows for all values of $s$. Similarly there are $n$ linearly independent columns if the matrix $C$ has full rank. This implies that systems where the matrices $B$ or $C$ are of full rank do not have zeros. In particular it means that a system has no zeros if the full state is measured.

## 6.3 Bode Plots

A useful representation of the frequency response was proposed by Bode who represented it by two curves, the gain curve and the phase curve. The gain curve gives gain $|G(i\omega)|$ as a function of $\omega$ and the phase curve phase $\arg G(i\omega)$ as a function of $\omega$. The curves are plotted as shown below with logarithmic scales for frequency and magnitude and linear scale for phase, see

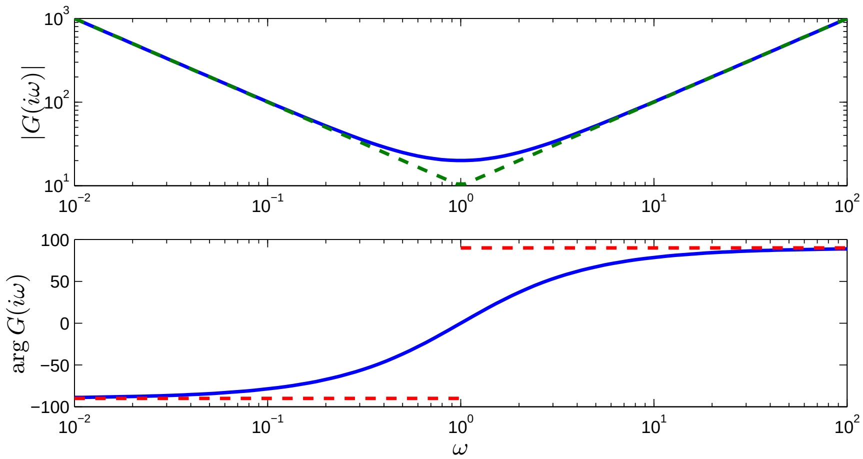


Figure 6.2: Bode plot of the transfer function of the ideal PID controller $C(s) = 20 + 10/s + 10s$. The top plot is the gain curve and bottom plot is the phase curve. The dashed lines show straight line approximations of the curves.

Figure 6.2 An useful feature of the Bode plot is that both the gain curve and the phase curve can be approximated by straight lines, see Figure 6.2 where the approximation is shown in dashed lines. This fact was particularly useful when computing tools were not easily accessible. The fact that logarithmic scales were used also simplified the plotting.

It is easy to sketch Bode plots because with the right scales they have linear asymptotes. This is useful in order to get a quick estimate of the behavior of a system. It is also a good way to check numerical calculations.

Consider first a transfer function which is a polynomial $G(s) = b(s)/a(s)$. We have

$$\log G(s) = \log b(s) - \log a(s)$$

Since a polynomial is a product of terms of the type :

$$s, \quad s + a, \quad s^2 + 2\zeta as + a^2$$

it suffices to be able to sketch Bode diagrams for these terms. The Bode plot of a complex system is then obtained by composition.

*Example* 18 *(Bode Plot of a Differentiator).* Consider the transfer function

$$G(s) = s$$

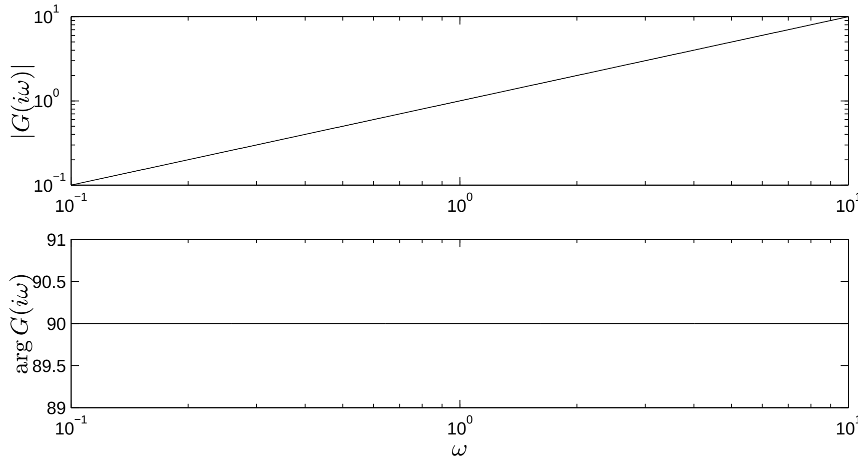


Figure 6.3: Bode plot of the transfer function $G(s) = s$, i.e. a differentiator.

We have $G(i\omega) = i\omega$ which implies

$$\log |G(i\omega)| = \log \omega$$
$$\arg G(i\omega) = \pi/2$$

The gain curve is thus a straight line with slope 1 and the phase curve is a constant at 90°.. The Bode plot is shown in Figure 6.3

*Example* 19 *(Bode Plot of an Integrator).* Consider the transfer function

$$G(s) = \frac{1}{s}$$

We have $G(i\omega) = 1/i\omega$ which implies

$$\log |G(i\omega)| = -\log \omega$$
$$\arg G(i\omega) = -\pi/2$$

The gain curve is thus a straight line with slope -1 and the phase curve is a constant at $-90°$. The Bode plot is shown in Figure 6.4

Compare the Bode plots for the differentiator in Figure 6.3 and the integrator in Figure 6.4. The sign of the phase is reversed and the gain curve is mirror imaged in the horizontal axis. This is a consequence of the property of the logarithm.

$$\log \frac{1}{G} = -\log G = -\log |G| - i \arg G$$

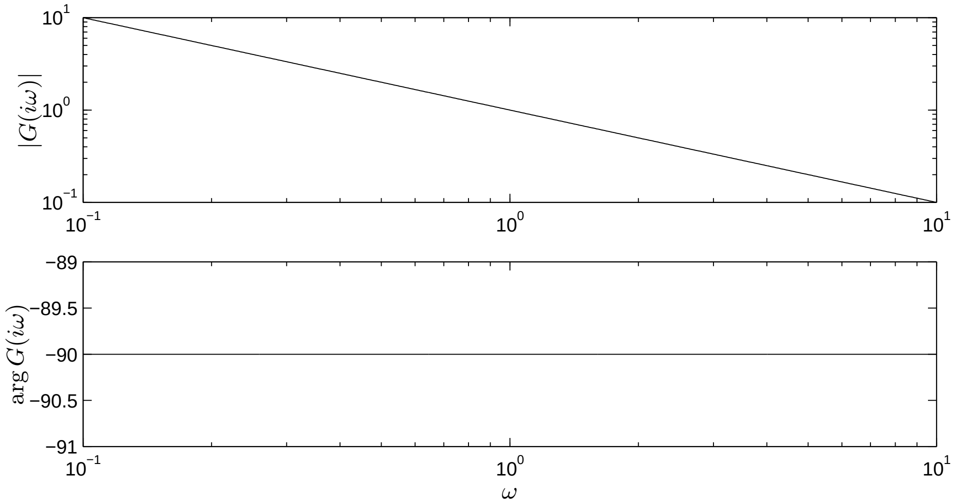


Figure 6.4: Bode plot of the transfer function $G(s) = 1/s$, i.e. an integrator.

*Example* 20 *(Bode Plot of a First Order Factor).* Consider the transfer function

$$G(s) = s + a$$

We have

$$G(i\omega) = a + i\omega$$

and it follows that

$$|G(i\omega)| = \sqrt{\omega^2 + a^2}, \quad \arg G(i\omega) = \arctan \omega/a$$

Hence

$$\log |G(i\omega)| = \frac{1}{2} \log (\omega^2 + a^2), \quad \arg G(i\omega) = \arctan \omega/a$$

The Bode Plot is shown in Figure 6.5. Both the gain curve and the phase curve can be approximated by straight lines if proper scales are chosen and

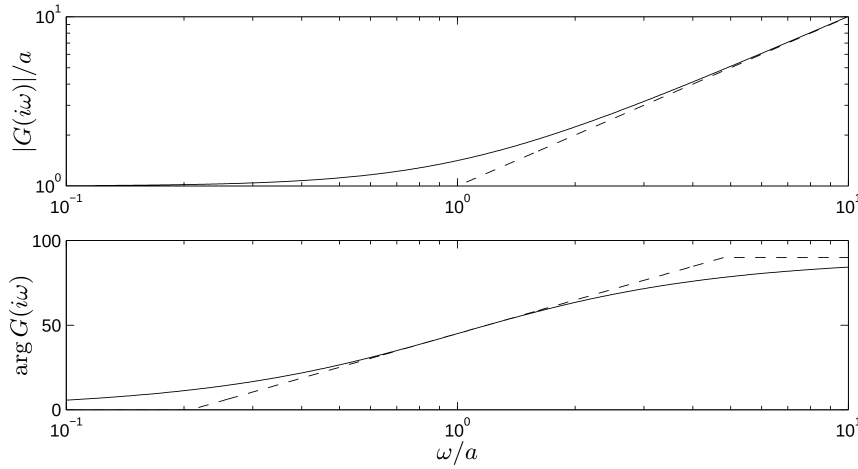


Figure 6.5: Bode plot of the transfer function $G(s) = s + a$. The dashed lines show the piece-wise linear approximations of the curves.

we obtain the following approximations.

$$\log|G(i\omega)| \approx \begin{cases} \log a & \textit{if } \omega << a, \\ \log a + \log\sqrt{2} & \textit{if } \omega = a, \\ \log\omega & \textit{if } \omega >> a \end{cases},$$

$$\arg G(i\omega) \approx \begin{cases} 0 & \textit{if } \omega << a, \\ \frac{\pi}{4} + \frac{1}{2}\log\frac{\omega}{a} & \textit{if } \omega \approx a, \\ \frac{\pi}{2} & \textit{if } \omega >> a \end{cases},$$

Notice that a first order system behaves like an integrator for high frequencies. Compare with the Bode plot in Figure 6.4.

*Example* 21 *(Bode Plot of a Second Order System).* Consider the transfer function

$$G(s) = s^2 + 2a\zeta s + a^2$$

We have

$$G(i\omega) = a^2 - \omega^2 + 2i\zeta a\omega$$

Hence

$$\log|G(i\omega)| = \frac{1}{2}\log\left(\omega^4 + 2a^2\omega^2(2\zeta^2 - 1) + a^4\right)$$
$$\arg G(i\omega) = \arctan 2\zeta a\omega/(a^2 - \omega^2)$$

Notice that the smallest value of the magnitude $\min_\omega|G(i\omega)| = 1/2\zeta$ is obtained for $\omega = a$ The gain is thus constant for small $\omega$. It has an asymptote with zero slope for low frequencies. For large values of $\omega$ the gain is proportional to $\omega^2$, which means that the gain curve has an asymptote with slope 2. The phase is zero for low frequencies and approaches 180° for large frequencies. The curves can be approximated with the following piece-wise linear expressions

$$\log|G(i\omega)| \approx \begin{cases} 2\log a & \textit{if } \omega << a, \\ 2\log a + \log 2\zeta & \textit{if } \omega = a, \\ 2\log\omega & \textit{if } \omega >> a \end{cases},$$

$$\arg G(i\omega) \approx \begin{cases} 0 & \textit{if } \omega << a, \\ \frac{\pi}{2} + \frac{\omega - a}{a\zeta} & \textit{if } \omega = a, \\ \pi & \textit{if } \omega >> a \end{cases},$$

The Bode Plot is shown in Figure 6.6, the piece-wise linear approximations are shown in dashed lines.

### Sketching a Bode Plot*

It is easy to sketch the asymptotes of the gain curves of a Bode plot. This is often done in order to get a quick overview of the frequency response. The following procedure can be used

- Factor the numerator and denominator of the transfer functions.
- The poles and zeros are called break points because they correspond to the points where the asymptotes change direction.
- Determine break points sort them in increasing frequency
- Start with low frequencies
- Draw the low frequency asymptote

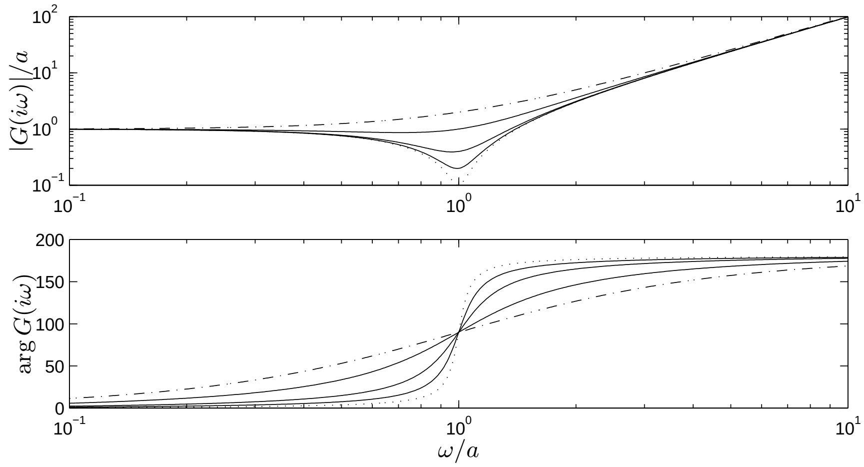


Figure 6.6: Bode plot of the transfer function $G(s) = s^2 + 2\zeta as + a^2$ with $\zeta = 0.05$ (dotted), 0.1, 0.2, 0.5 and 1.0 (dash-dotted). The dashed lines show the piece-wise linear approximations of the curves.

- Go over all break points and note the slope changes
- A crude sketch of the phase curve is obtained by using the relation that, for systems with no RHP poles or zeros, one unit slope corresponds to a phase of 90°

We illustrate the procedure with the transfer function

$$G(s) = \frac{200(s+1)}{s(s+10)(s+200)} = \frac{1+s}{10s(1+0.1s)(1+0.01s)}$$

The break points are 0.01, 0.1, 1. For low frequencies the transfer function can be approximated by

$$G(s) \approx \frac{1}{10s}$$

Following the procedure we get

- The low frequencies the system behaves like an integrator with gain 0.1. The low frequency asymptote thus has slope -1 and it crosses the axis of unit gain at $\omega = 0.1$.
- The first break point occurs at $\omega = 0.01$. This break point corresponds to a pole which means that the slope decreases by one unit to -2 at that frequency.

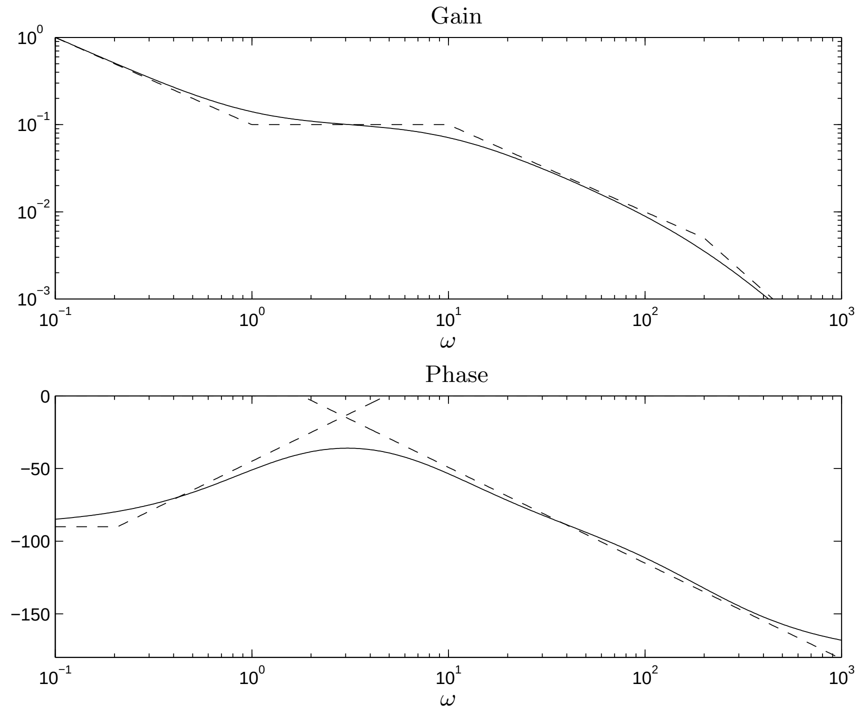


Figure 6.7: Illustrates how the asymptotes of the gain curve of the Bode plot can be sketched. The dashed curves show the asymptotes and the full lines the complete plot.

- The next break point is at $\omega = 0.1$ this is also a pole which means that the slope decreases to -3.
- The next break point is at $\omega = 1$, since this is a zero the slope increases by one unit to -2.

Figure 6.7 shows the asymptotes of the gain curve and the complete Bode plot.

## Interpretations of Bode Plots

The Bode plot gives a quick overview of the properties of a system. Many properties can be read off directly from the plot. The plots give primarily the gain and the phase for different frequencies. Since it is possible to decompose any signal into a sum of sinusoids it is possible to visualize the behavior of a system for different frequency ranges. Furthermore when the gain curves are close to the asymptotes the system can be approximated by integrators of differentiators. Consider for example the Bode plot in Figure 6.2. For low frequencies the gain curve of the Bode plot has the slope -1 which means that the system acts like an integrator. For high frequencies the gain curve has slope +1 which means that the system acts like a differentiator for high frequencies.

## 6.4 Block Diagrams

Feedback systems are often large and complex. It is therefore a major challenge to understand, analyze and design them. This is illustrated by the fact that the idea of feedback was developed independently in many different application areas. It took a long time before it was found that the systems were based on the same idea. The similarities became apparent when proper abstractions were made. In this section we will develop some ideas that are used to describe feedback systems. The descriptions we are looking for should capture the essential features of the systems and hide unnecessary details. They should be applicable to many different systems.

### Schematic Diagrams

In all branches of engineering, it is common practice to use some graphical description of systems. They can range from stylistic pictures to drastically simplified standard symbols. These pictures make it possible to get an overall view of the system and to identify the physical components. Examples of such diagrams are shown in Figure 6.8

### Block Diagrams

The schematic diagrams are useful because they give an overall picture of a system. They show the different physical processes and their interconnection, and they indicate variables that can be manipulated and signals that can be measured.

A special graphical representation called *block diagrams* has been developed in control engineering. The purpose of block diagrams is to emphasize the information flow and to hide technological details of the system. It is natural to look for such representations in control because of its multidisciplinary nature. In a block diagram, different process elements are shown as boxes. Each box has inputs denoted by lines with arrows pointing toward

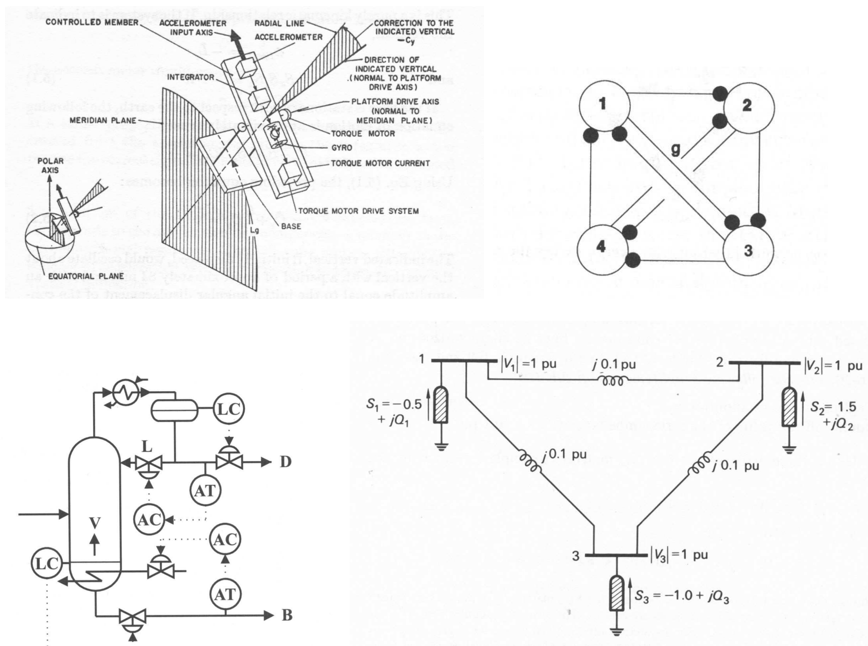


Figure 6.8: Examples of schematic descriptions: a schematic picture of an inertial navigation system (upper left), a neuron network for respiratory control (upper right), a process and instrumentation diagram (lower left) and a power system (lower right).

the box and outputs denoted by lines with arrows going out of the box. The inputs denote the variables that influence a process and the outputs denote some consequences of the inputs that are relevant to the feedback system.

Figure 6.9 illustrates how the principle of information hiding is used to derive an abstract representation of a system. The upper part of the picture shows a photo of a physical system which a small desk-top process in a control laboratory. It consists of two tanks, a pump that pumps water to the tanks, sensors, and a computer which implements the control algorithm and provides the user interface. The purpose of the system is to maintain a specified level in the lower tank. To do so, it is necessary to measure the level. The level can be influenced by changing the speed of the motor that pumps water into the upper tank. The voltage to the amplifier that drives the pump is selected as the control variable. The controller receives information about the desired level in the tank and the actual tank level. This is accomplished using an AD converter to convert the analog signal

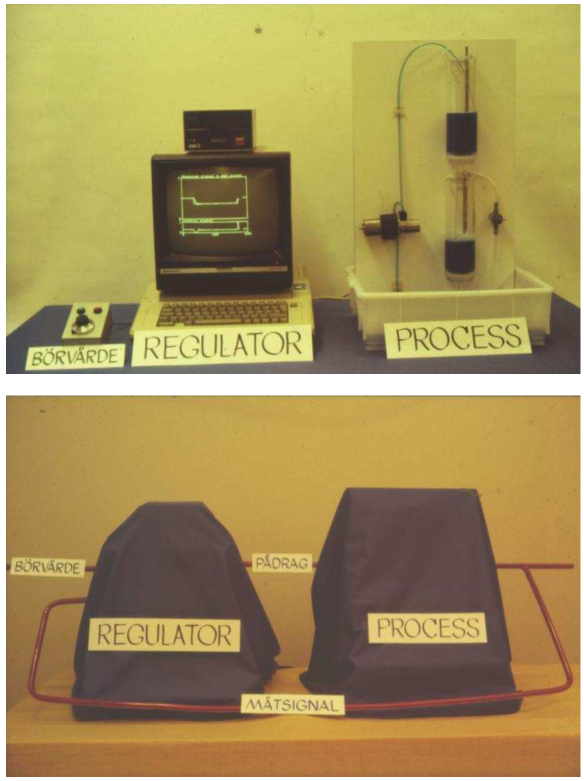


Figure 6.9: Illustrates the process of information hiding used to obtain a block diagram. The top figure is a picture of the physical system, the middle figure is obtained by hiding many details about the system and the bottom figure is the block diagram.

to a number in the computer. The control algorithm in the computer then computes a numerical value of the control variable. This is converted to a voltage using a DA converter. The DA converter is connected to an amplifier for the motor that drives the pump.

The first step in making a block diagram is to identify the important signals: the control variable, the measured signals, disturbances and goals. Information hiding is illustrated in the figure by covering systems by a cloth as shown in the lower part of Figure 6.9. The block diagram is simply a stylized picture of the systems hidden by the cloth.

In Figure 6.9, we have chosen to represent the system by two blocks only. This granularity is often sufficient. It is easy to show more details

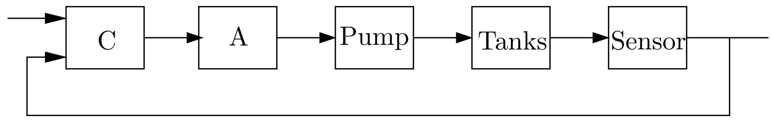


Figure 6.10: A more detailed block diagram of the system in Figure 6.9 showing controller $C$, amplifier $A$, pump, tanks and sensor.

simply by introducing more subsystems, as indicated in Figure 6.10 where we show the drive amplifier, motor, pump, and tanks, the sensors with electronics, the AD converter, the computer and the DA converter. The detail chosen depends on the aspects of the system we are interested in and the taste of the person doing the investigation. Remember that parsimony is a trademark of good engineering. Very powerful tools for design, analysis and simulation were developed when the block diagrams were complemented with descriptions of the blocks in terms of transfer functions.

### Causality

The arrows in a block diagram indicate causality because the output of a block is caused by the input. To use the block diagram representation, it is therefore necessary that a system can be partitioned into subsystems with causal dependence. Great care must be exercised when using block diagrams for detailed physical modeling as is illustrated in Figure 6.11. The tank system in Figure 6.11B is a cascade combination of the two tanks shown in Figure 6.11B. It cannot be represented by cascading the block diagram representations because the level in the second tank influences the flow between the tanks and thus also the level in the first tank. When using block diagrams it is therefore necessary to choose blocks to represent units which can be represented by causal interactions. We can thus conclude that even if block diagrams are useful for control they also have serious limitation. In particular they are not useful for serious physical modeling which has to be dealt with by other tools which permit bidirectional connections.

### Examples

An important consequence of using block diagrams is that they clearly show that control systems from widely different domains have common features because their block diagrams are identical. This observation was one of the key factors that contributed to the emergence of the discipline of automatic

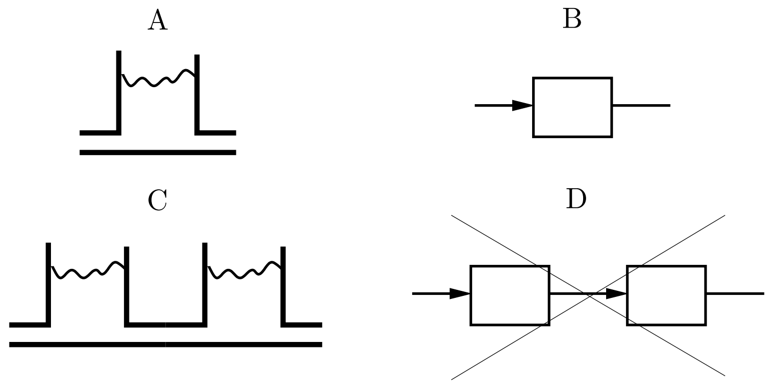


Figure 6.11: A simple hydraulic system with an inflow and a free outflow is shown in A. The block diagram representation of the system is shown in B. The system obtained by connecting two hydraulic systems is shown in C. This system cannot be represented by the series connection of the block diagrams in B.

control in the 1940s. We will illustrate this by showing the block diagrams of some of the systems discussed in Chapter **??**.

*Example* 22 *(A steam engine with a centrifugal governor).* The steam engine with the centrifugal governor in Example 22 can be represented with the block diagram shown in Figure 6.12. In this block diagram we have chosen to represent the steam engine with one block. This block has two inputs: the position of the steam valve and the load torque of the systems that the engine is driving. The system has one output which is engine speed. The controller is a box with two inputs: the engine speed and desired engine speed. The output of the controller is the steam valve position. There is some two-way interaction between the controller and the valve position but with appropriate gearing and heavy balls in the governor it may be assumed that the force exerted by the valve on the governor is negligible.

*Example* 23 *(An aircraft stabilizer).* To develop a block diagram for an airplane with the Sperry stabilizer, we first introduce suitable variables and describe the system briefly. The pitch angle that describes how the airplane is pointing is an important variable and is measured by the gyro-stabilized pendulum. The pitch angle is influenced by changing the rudder. We choose to represent the airplane by one box whose input is the rudder angle and whose output is the pitch angle. There is another input representing the forces on the airplane from wind gusts. The stabilizer attempts to keep the

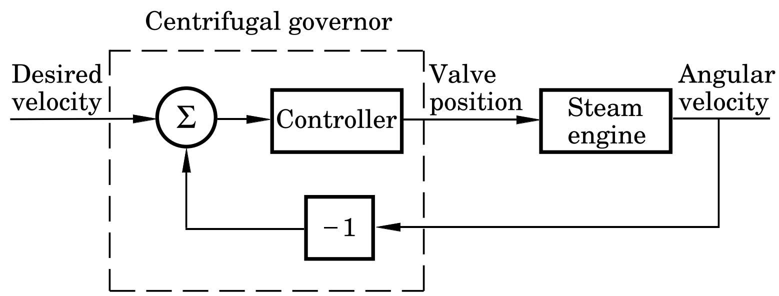


Figure 6.12: Block diagram of a steam engine with a centrifugal governor.

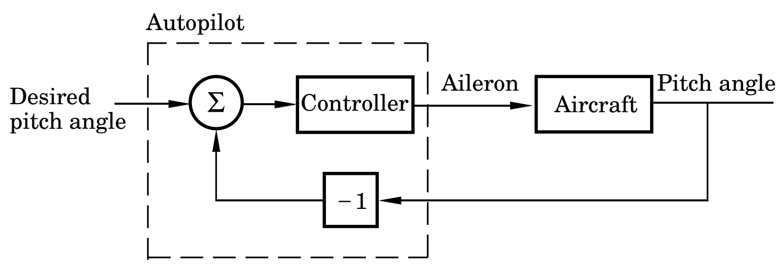


Figure 6.13: Block diagram of an airplane with the Sperry autopilot.

pitch angle small by appropriate changes in the rudder. This is accomplished by wires that connect the rudder to the gyro-stabilized pendulum. There is also a mechanism enabling the pilot to choose a desired value of the pitch angle if he wants the airplane to ascend or descend. In the block diagram we represent the controller with one block where the difference between desired and actual pitch angles is the input and the rudder angle is the output. Figure 6.13 shows the block diagram obtained.

Even if block diagrams are simple, it is not always entirely trivial to obtain them. It happens frequently that individual physical components to not necessarily correspond to specific blocks and that it may be necessary to use mathematics to obtain the block. We illustrate this by an example.

*Example* 24 *(A feedback amplifier).* An electronic amplifier with negative feedback was discussed in Section **??**. A schematic diagram of the amplifier is shown in Figure 6.14. To develop a block diagram we first decide to represent the pure amplifier as one block. This has input $V$ and output $V_2$.

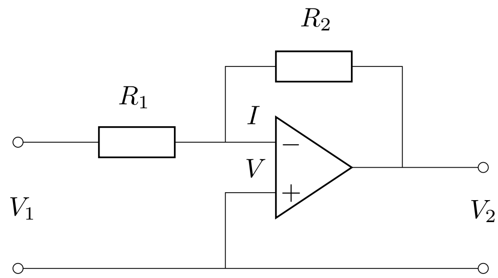


Figure 6.14: A feedback amplifier.

The input-output relation is

$$V_2 = -GV$$

where $G$ is the gain of the amplifier and the negative sign indicates negative feedback. If the current $I$ into the amplifier is negligible the current through resistors $R_1$ and $R_2$ are the same and we get

$$\frac{V_1 - V}{R_1} = \frac{V - V_2}{R_2}$$

Solving this equation for the input voltage $V$ to the amplifier we get

$$V = \frac{R_2 V_1 + R_1 V_2}{R_1 + R_2} = \frac{R_2}{R_1 + R_2}\Big(V_1 + \frac{R_1}{R_2} V_2\Big)$$

This equation can be represented by one block with gain $R_2/(R_1 + R_2)$ and the input $V_1 + R_1 V_2/R_1$ and we obtain the block diagram shown in Figure 6.15. The lower representation where the process has positive gain and the feedback gain is negative has become the standard of representing feedback systems.

Notice that the individual resistors do not appear as individual blocks, they actually appear in various combinations in different blocks. This is one of the difficulties in drawing block diagrams. Also notice that the diagrams can be drawn in many different ways. The middle diagram in Figure 6.15 is obtained by viewing $-V_2$ as the output of the amplifier. This is the standard convention where the process gain is positive and the feedback gain is negative. The lowest diagram in Figure 6.15 is yet another version, where the ratio $R_1/R_2$ is brought outside the loop. In all three diagrams the gain around the loop is $R_2 G/(R_1 + R_2)$, this is one of the invariants of a feedback system.

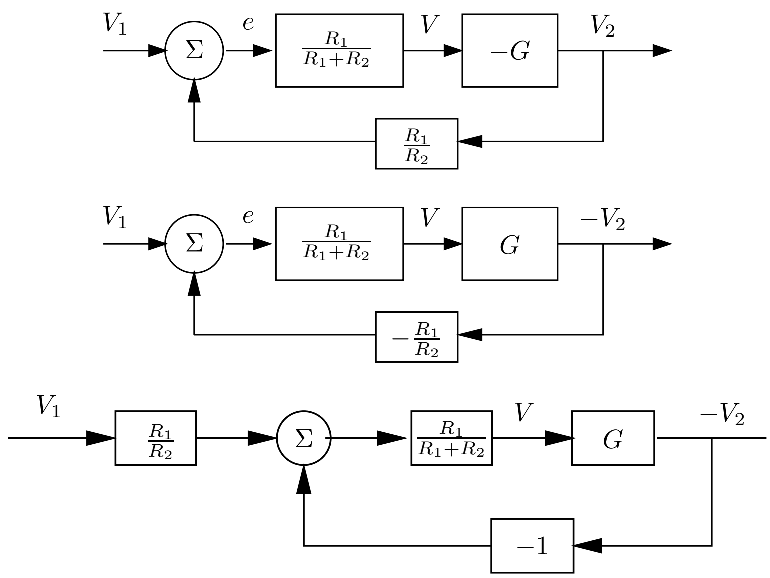


Figure 6.15: Three block diagrams of the feedback amplifier in Figure 6.14.

### A Generic Control System with Error Feedback

Although the centrifugal governor, the autopilot and the feedback amplifier in Examples 22, 23 and 24 represent very different physical systems, their block diagrams are identical apart from the labeling of blocks and signals, compare Figures 6.12, 6.13 and 6.15. This illustrates the universality of control. A generic representation of the systems is shown in Figure 6.16. The system has two blocks. One block $P$ represents the process and the other $C$ represents the controller. Notice negative sign of the feedback. The signal $r$ is the reference signal which represents the desired behavior of the process variable $x$.

Disturbances are an important aspect of control systems. In fact if there were no disturbances there is no reason to use feedback. In Figure 6.16 there are two types of disturbances, labeled $d$ and $n$. The disturbance labeled $d$ is called a load disturbance and the disturbance labeled $n$ is called measurement noise. Load disturbances drive the system away from its desired behavior. In Figure 6.16 it is assumed that there is only one disturbance that enters at the system input. This is called an input disturbance. In practice there may be many different disturbances that enter the system in many different ways. Measurement noise corrupts the information about

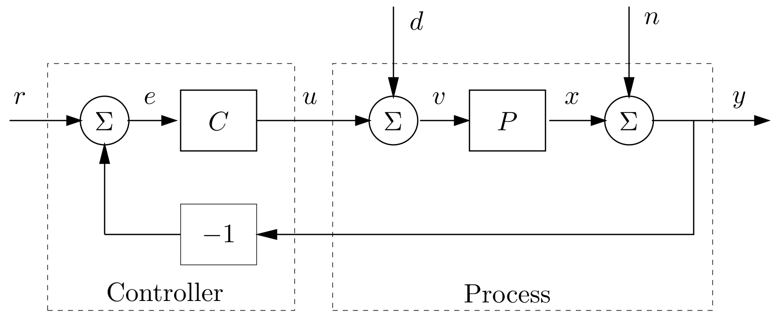


Figure 6.16: Generic control system with error feedback.

the process variable obtained from the measurements. In Figure 6.16 it is assumed that the measured signal $y$ is the sum of the process variable $x$ and measurement noise. In practice the measurement noise may appear in many other ways.

The system in Figure 6.16 is said to have error feedback, because the control actions are based on the error which is the difference between the reference $r$ and the output $y$. In some cases like a CD player there is no explicit information about the reference signal because the only information available is the error signal. In such case the system shown in Figure 6.16 is the only possibility but if the reference signal is available there are other alternatives that may give better performance.

### Block Diagrams and Transfer Functions

Transfer functions are very well suited for analysis of linear control systems through the use of *block diagram alegbra*. Given a system with transfer function $H(s)$, we can symbolically represent the input/output response as

$$\frac{Y(s)}{U(s)} = H(s).$$

Here $U(s)$ and $Y(s)$ formally represent the Laplace transforms of the signals $u$ and $y$, but for now we can simply think of these as placeholders for the input and output of the system, respectively. We can thus write

$$Y(s) = H(s)U(s).$$

Using this notation, a block can thus simply be represented algebraically as a multiplication. Since the other element of a block diagram is a summation it follows that relations between signals in a block diagram can be

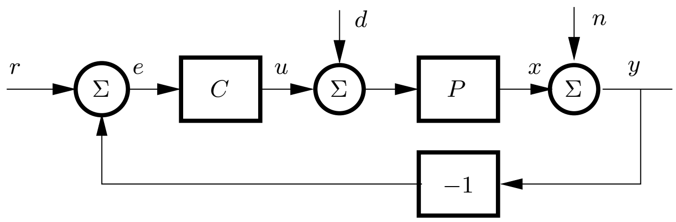


Figure 6.17: Block diagram of a feedback system.

obtained by pure algebraic manipulations. We thus obtain the following very simple recipe for analyzing control systems.

- Draw the block digram of the system.
- Derive transfer functions for each block.
- Use algebra to obtain the transfer functions that relate the signals of interest.
- Interprete the transfer function.
- Simulate the system by computing responses to interesting signals.

The combination of block diagrams and transfer functions is a powerful because it is possible both to obtain an overview of a system and find details of the behavior of the system. By representing signals by their Laplace transforms and the blocks by transfer functions the relations between signals in the system are obtained by straight forward manipulations. We illustrate this by an example.

*Example* 25 *(Relations Between Signals in a Block Diagram).* Consider the system in Figure 6.17. The system has two blocks representing the process $P$ and the controller $C$. There are three external signals, the reference $r$, the load disturbance $d$ and the measurement noise $n$. A typical problem is to find out how the error $e$ related to the signals $r$ $d$ and $n$? Introduce the Laplace transforms of the signals and the transfer functions of the blocks. To simplify the notation we will drop the argument $s$ of the Laplace transforms. Signals are labeled by lower case letters and their Laplace transforms with the corresponding upper case letters.

To obtain the desired relation we simply trace the signals around the loop. We begin with the signal we are interested in, i.e. $e$. It follows from the block diagram that the Laplace transform of the error $E$ is given by

$$E = R - Y.$$

The signal $y$ in turn is the sum of $n$ and $x$, hence

$$Y = N + X,$$

where $x$ is the output of the process, i.e.

$$X = PV = P(D + U),$$

where $u$ is the output of the controller, i.e.

$$U = CE.$$

Combining the equations gives

$$\begin{aligned} E &= R - Y = R - (N + X) = R - \big(N + P(D + U)\big) \\ &= R - \big(N + P(D + CE)\big). \end{aligned}$$

Hence

$$E = R - \big(N + P(D + CE)\big) = R - N - PD - PCE. \tag{6.6}$$

Solving this equation for $E$ gives

$$E = \frac{1}{1 + PC} R - \frac{1}{1 + PC} N - \frac{P}{1 + PC} D$$

Hence

$$E = \frac{1}{1 + PC} R - \frac{1}{1 + PC} N - \frac{P}{1 + PC} D = G_{er} R + G_{en} N + G_{ed} D \tag{6.7}$$

The error is thus the sum of three terms, depending on the reference $r$, the measurement noise $n$ and the load disturbance $d$. The function

$$G_{er} = \frac{1}{1 + PC}$$

is the transfer function from reference $r$ to error $e$, $G_{en}$ is the transfer functions from measurement noise $n$ to error $e$ and $G_{ed}$ is the transfer functions from load disturbance $d$ to error $e$. It follows from (6.7) that the Laplace transform of the error is a sum of three terms where each term is a product of a transfer function and a signal. This is a typical illustration of the superposition principle.

The example illustrates an effective way to manipulate the equations to obtain the relations between inputs and outputs in a feedback system. The general idea is to start with the signal of interest and to trace signals around the feedback loop until coming back to the signal we started with. With a some practice the equation (6.6) can be written directly by inspection of the block diagram. Notice that all terms in Equation (6.7) formally have the same denominators, there may, however, be factors that cancel.

## 6.5 Further Reading