Introduction to Control System Modeling

1 Introduction

A **mathematical model** is an equation or set of equations which adequately describes the behavior of a system.

Control system modeling is a subject in its own right. Essentially there are two approaches to finding the model.

- In the first, the system is broken down into smaller elements. For each element a mathematical description is then established by working from the physical laws which describe the system's behavior. The simplest such technique is lumped-parameter modeling, which is considered in this course.
- The second approach is known as system identification in which it is assumed that an experiment can be carried out on the system, and that a mathematical model can be found from the results. This approach can clearly only be applied to existing plants, whereas lumped-parameter modeling can additionally be applied to a plant yet to be built, working purely from the physics of the proposed plant components.

Control system modeling is a specialization of the more general area of mathematical modeling. In a control system model, the important relationship is that between the manipulated inputs and measurable outputs. Ideally this relationship should be linear and capable of being described by an expression of low-order (that is, an equation or set of equations containing as few terms as possible). An usual graphical representation of single input, single output (SISO) dynamical system is shown in Figure 1. A model of such a system is a mathematical relationship between the input $u(t)$ and the output $y(t)$.

Figure 1: SISO dynamic system

Most commonly, a low-order linear differential equation model is used.

When fundamental physical laws are applied to the lumped-parameter model, the resulting equations may be nonlinear in which case further assumptions may have to be made in order to produce an ordinary linear differential equation model which is soluble. In such cases, it is not unusual to assume that system operation will be restricted to small perturbations about a given operating condition. If the assumed operating region is small enough, most nonlinear plants may be described by a set of linear equations. If these assumptions cannot be made, nonlinear control techniques may be appropriate.

2 Lumped-parameter models

The systems studied in this course are:

Linear - It must obey the principle of superposition that is, if an input $u_1(t)$ causes an output $y_1(t)$ and an input $u_2(t)$ causes an output $y_2(t)$, then an input $u_1(t)+u_2(t)$ causes an output $y_1(t)+y_2(t)$ (see next section **Linear Approximations of Physical systems**)

In the case of the resistor, if an input of 1A produces an output of 3V, and an input of 2A produces an output of 6V then an input of 3A produces an output of 9V.

Stationary (or time invariant) - The parameters inside the element must not vary with time. In other words, an input applied today must give the same result as the same input applied yesterday or tomorrow.

A vehicle which burns large masses of fuel, such as racing cars or space vehicles, is an example of system which is **not** stationary. Its dynamic behavior will alter significantly as its mass decreases.

Deterministic - The outputs of the system at any time can be determined from a knowledge of the system's inputs up to that time. In other words there is no random (or stochastic) behavior in the system, since its outputs are always a specific function of the inputs. The term *causal* is also used for such systems.

Examples. Consider an element representing an idealized component such as:

- The resistor. If the resistor is linear then its model is a relationship between the voltage $v(t)$ and the current $i(t)$:

$$i(t) = \frac{1}{R}v(t)$$

- The inductor is described by the equation:

$$i(t) = \frac{1}{L}\int v(t)dt \quad \text{or} \quad v(t) = L\frac{di(t)}{dt}$$

- The capacitor is described by:

$$i(t) = C\frac{dv(t)}{dt}$$

- The spring-mass-damper mechanical system (Figure 2). Assuming the system is linear, a model of the system can be derived from the Newton's second law of motion. It is a relationship between the (input) force $r(t)$ and the (output) displacement $y(t)$. The system could represent for example an automobile shock absorber. The

$$M\frac{d^2y(t)}{dt^2} + f\frac{dy(t)}{dt} + ky(t) = r(t)$$

where: f is the friction coefficient, M - the mass, k - the stiffness of the linear spring.

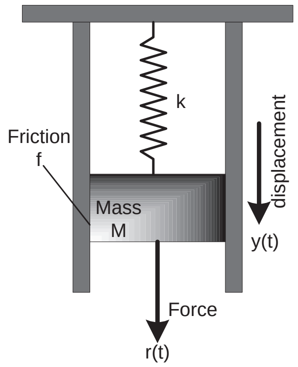

Figure 2: Spring-mass-damper system

3 Linear Approximation of Physical Systems

A great majority of physical systems are linear within some range of variables. However, all systems ultimately become nonlinear as the variables are increased without limit.

A system is defined as linear in terms of the system excitation and response. In the case of an electrical network, the excitation is the input current and the response is the voltage. In general a *necessary condition* for a linear system can be determined in terms of and excitation $x(t)$ and a response $y(t)$. When a system at rest is subjected to an excitation $x_1(t)$, it provides a response $y_1(t)$. Furthermore, when the system is subjected to an excitation $x_2(t)$, it provides a corresponding response $y_2(t)$. For a linear system, it is *necessary* that the excitation $x_1(t)+x_2(t)$ result in a response $y_1(t) + y_2(t)$. This is usually called the *principle of superposition.*

Furthermore, it is necessary that the magnitude scale factor be preserved in a linear system. Again, consider a system with an input x and an output y. Then it is necessary that the response of a linear system to a constant multiple β of an input x be equal to the response to the input multiplied by the same constant so that the output is equal to βy. This is called the property of *homogeneity.*

Example 1. A system characterized by the relation $y = x^2$ is not linear because the superposition property is not satisfied. A system represented by the relation $y = mx + b$ is not linear because it does not satisfy the homogeneity property. However, this device may be considered linear about an operating point x_0, y_0 for small changes Δx and Δy. When $x = x_0 + \Delta x$ and $y = y_0 + \Delta y$, we have

$$y = mx + b$$

or

$$y_0 + \Delta y = mx_0 + m\Delta x + b$$

and therefore

$$\Delta y = m\Delta x$$

which satisfies the necessary conditions.

Consider a general element with an excitation $x(t)$ and a response $y(t)$. The relationship of the two variables is written as:

$$y(t) = g(x(t))$$

where $g(x(t))$ indicates $y(t)$ is a function of $x(t)$. The normal operating point is designated by x_0. Because the curve (function) is continuous over the range of interest, a Taylor series expansion about the operating point may be utilized. Then we have:

$$y = g(x) = g(x_0) + \frac{dg}{dx}|_{x=x_0}\frac{x - x_0}{1!} + \frac{d^2g}{dx^2}|_{x=x_0}\frac{(x-x_0)^2}{2!} + ...$$

The slope at the operating point,

$$\frac{dg}{dx}|_{x=x_0},$$

is a good approximation to the curve over a small range of $(x - x_0)$, the deviation from the operating point. Then, as a reasonable approximation, the expression from example 1 becomes:

$$y = g(x_0) + \frac{dg}{dx}|_{x=x_0}(x - x_0) = y_0 + m(x - x_0),$$

where m is the slope at the operating point. Finally, this equation can be rewritten as the linear equation

$$(y - y_0) = m(x - x_0)$$

or

$$\Delta y = m\Delta x.$$

If the dependent variable y depends upon several excitation variables $x_1, x_2, ..., x_n$, then the functional relationship is written as:

$$y = g(x_1, x_2, ..., x_n).$$

The Taylor series expansion about the operating point $x_{10}, x_{20}, ..., x_{n0}$ is useful for a linear approximation to the nonlinear function. When the higher-order terms are neglected, the linear approximation is written as

$$y = g(x_{10}, x_{20}, ..., x_{n0}) + \frac{dg}{dx_1}|_{x=x_0}(x_1 - x_{10}) + \frac{dg}{dx_2}|_{x=x_0}(x_2 - x_{20}) + ... + \frac{dg}{dx_n}|_{x=x_0}(x_n - x_{n0})$$

where x_0 is the operating point. An example will clearly illustrate this method.

Example 2. Pendulum oscillator model.

Consider the pendulum oscillator shown in Figure 3. The torque on the mass is

$$T = MgLsin(x)$$

where g is the gravity constant. The equilibrium condition for the mass is $x_0 = 0^o$. The first derivative evaluated at equilibrium provides the linear approximation, which is:

$$T - T_0 \cong MgL\frac{\partial sinx}{\partial x}|_{x=x_0}(x - x_0),$$

where $T_0 = 0$. Then, we have

$$T = MgL(cos0^o)(x - 0^o) = MgLx$$

The approximation is reasonably accurate for $-\pi/4 \leq x \leq \pi/4$. For example, the response of the linear model for the swing through $\pm 30^o$ is within 2% of the actual nonlinear pendulum response.

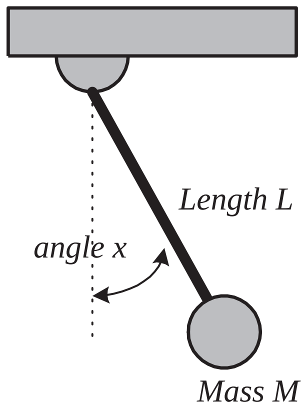

Figure 3: Pendulum oscillator

4 The Laplace Transform

The ability to obtain linear approximations of physical systems allows the analyst to consider the use of Laplace transformation. The Laplace transform method substitutes the relatively easily solved algebraic equations for the more difficult differential equations. The basic Laplace transform of a time signal f(t) is defined as:

$$F(s) = \int_0^\infty f(t)e^{-st}dt$$

and written symbolically as: $F(s) \equiv \pounds[f(t)]$

It is common practice to use a capital letter F which is a function of the new variable s for the transform of the time signal $f(t)$. Also it is assumed that $f(t)$ is zero for all times before t=0. In the definition of Laplace transform the exponent st must be dimensionless otherwise the expression e^{-st} is meaningless. Thus the variable s has dimension of 1/time which is the dimension of frequency. Since s is also a complex quantity it is often referred to as the *the complex frequency.*

5 Signals

A few commonly used signals will be introduced in this section:

1. The unit step is defined by (Figure 4 a):

$$u(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases} \tag{1}$$

or, if the signal is shifted in time (Figure 4 b):

$$u(t-\tau) = \begin{cases} 0, & t-\tau < 0 \\ 1, & t-\tau \geq 0 \end{cases} \tag{2}$$

Table 1: Laplace transform operations

1	Transform integral	f(t)	$\int_0^\infty f(t)e^{-st}dt$ or $\pounds[f(t)]$
2	Linearity	$f_1(t) \pm f_2(t)$	$F_1(s) \pm F_2(s)$
3	Constant multiplication	af(t)	aF(s)
4	Complex shift theorem	$e^{\pm at}f(t)$	F(s±a)
5	Real shift theorem	f(t-T)	$e^{-Ts}F(s)$, T≥0
6	Scaling theorem	f($\frac{t}{a}$)	aF(as)
7	First derivative	$\frac{d}{dt}$f(t)	sF(s) - f(0)
8	n-th derivative	$\frac{d^n}{dt^n}f(t)$	$s^nF(s) - \sum_{r=1}^{n} \frac{d^{r-1}}{dt^{r-1}}f(0)s^{n-r}$
9	First integral	$\int_0^t f(t)dt$	$\frac{1}{s}F(s)$

Table 2: Laplace transforms of common functions

	f(t)	F(s)
1	Unit impulse (Dirac) δ (t)	1
2	Unit step u(t)=1	$\frac{1}{s}$
3	Unit ramp v(t)=t	$\frac{1}{s^2}$
4	e^{at}	$\frac{1}{s-a}$
5	$cos\omega t$	$\frac{s}{s^2+\omega^2}$
6	$sin\omega t$	$\frac{\omega}{s^2+\omega^2}$

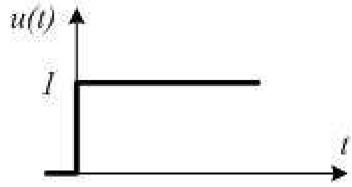

Figure 4: Unit step signal

The Laplace transform of the step function, according to Table 1 and 2, is:

$$\pounds[u(t)] = \frac{1}{s}, \quad \pounds[u(t-\tau)] = e^{-s\tau}\frac{1}{s} \tag{3}$$

2. The unit ramp is shown in Figure 5. It is defined by:

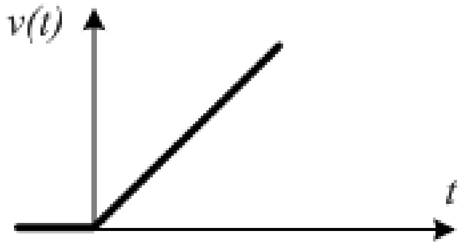

Figure 5: The ramp signal

$$v(t-\tau) = \begin{cases} 0, & t-\tau < 0 \\ t, & t-\tau \geq 0 \end{cases} \text{, or, for } \tau = 0, \quad v(t) = \begin{cases} 0, & t < 0 \\ t, & t \geq 0 \end{cases} \tag{4}$$

The Laplace transform of the ramp signal is:

$$\pounds[v(t)] = \frac{1}{s^2}, \quad \pounds[v(t-\tau)] = e^{-s\tau}\frac{1}{s^2} \tag{5}$$

3. The ideal impulse also called Dirac impulse is shown in Figure 6 It is defined by:

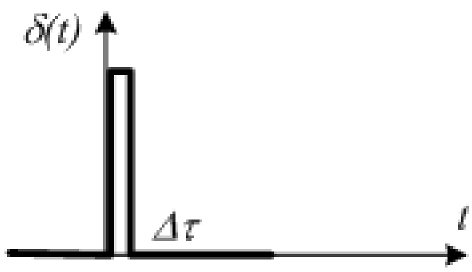

Figure 6: The impulse signal

$$\delta(t) = \begin{cases} 0, & t < \tau \;\; and \;\; t > \tau + \Delta\tau \\ A, & \tau \leq t \leq \tau + \Delta\tau \end{cases} \tag{6}$$

where

$$\lim_{\Delta\tau \to 0} \int_{\tau}^{\tau+\Delta\tau} \delta(t-\tau)dt = 1 \text{ and } A \to \infty \tag{7}$$

The Dirac impulse is a signal with unit area and infinite amplitude.

The Laplace transform of the unit impulse is:

$$\mathcal{L}[\delta(t-\tau)] = e^{-s\tau} \tag{8}$$

If the impulse is located at the origin:

$$\mathcal{L}[\delta(t)] = 1 \tag{9}$$

6 Transfer function models

The **transfer function** of a linear system is defined as the ratio of the Laplace transform of the output variable to the Laplace transform of the input variable, with all the initial conditions assumed to be zero.

A transfer function may be defined only for a linear stationary system. A transfer function is an input-output description of the behavior of a system.

A linear differential equation that describes such a system can be:

$$a_0 r(t) + a_1 \frac{dr(t)}{dt} + ... + a_m \frac{d^m r(t)}{dt^m} = b_0 y(t) + b_1 \frac{dy(t)}{dt} + ... + b_n \frac{d^n y(t)}{dt^n}$$

where $r(t)$ and $y(t)$ are the input and output variables, as shown in Figure 7. Applying the

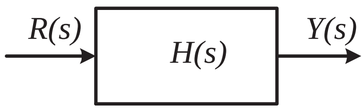

Figure 7: Block diagram of a system system

Laplace transform for the initial conditions zero, the differential equation becomes:

$$(a_0 + a_1 s + ... + a_m s^m)R(s) = (b_0 + b_1 s + ... + b_n s^n)Y(s)$$

and the transfer function will then be:

$$H(s) = \frac{Y(s)}{R(s)} = \frac{a_0 + a_1 s + ... + a_m s^m}{b_0 + b_1 s + ... + b_n s^n}$$

Example. Spring-mass-damper.The transfer function of the spring-mass-damper system, shown in Figure 2, is obtained from the original equation model by applying the Laplace transform:

$$M \frac{d^2 y(t)}{dt} + f \frac{dy(t)}{dt} + ky(t) = r(t)$$

$$Ms^2Y(s) + fsY(s) + kY(s) = R(s)$$

$$H(s) = \frac{Y(s)}{R(s)} = \frac{1}{Ms^2 + fs + k}$$

The block diagram now contains all the information given in the transfer function model, that is, "output = contents x input". $H(s)$ is called a transfer function. It shows how the input is transferred to the output.

Example. Electrical system. The transfer function for the system shown in Figure 8 can be derived as follows: Write the fundamental relation for each element:

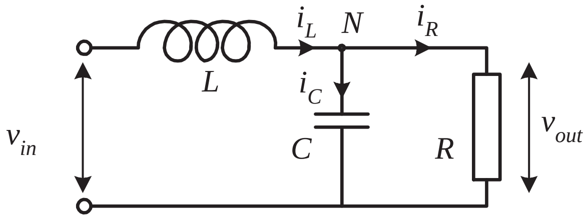

Figure 8: Electrical system

$$\text{Inductor:} \quad \frac{di_L}{dt} = \frac{1}{L} v_L \tag{10}$$

$$\text{Capacitor:} \quad \frac{dv_C}{dt} = \frac{1}{C} i_C \tag{11}$$

$$\text{Resistor:} \quad v_R = Ri_R \tag{12}$$

Write the Kirchhoff's current law at node N:

$$i_L = i_C + i_R \tag{13}$$

Write the Kirchhoff's voltage law:

$$v_{in} = v_L + v_C \tag{14}$$

$$v_C = v_R = v_{out} \tag{15}$$

Assume the initial conditions zero, replace (15) and apply the Laplace transform to the relations (10) - (14):

$$sI_L(s) = \frac{1}{L} V_L(s) \tag{16}$$

$$sV_{out}(s) = \frac{1}{C} I_C(s) \tag{17}$$

$$V_{out}(s) = RI_R(s) \tag{18}$$

$$I_L(s) = I_C(s) + I_R(s) \tag{19}$$

$$V_{in}(s) = V_L(s) + V_{out}(s) \tag{20}$$

Calculate to eliminate I_L, I_C, I_R, V_L from (16) - (20):

$$V_{in} = sLI_L + V_{out} = sL(I_C + I_R) + V_{out} = sL(sCV_{out} + \frac{1}{R}V_{out}) + V_{out} = (LCs^2 + \frac{L}{R}s + 1)V_{out} \quad (21)$$

The transfer function is:

$$H(s) = \frac{V_{out}(s)}{V_{in}(s)} = \frac{1}{LCs^2 + \frac{L}{R}s + 1} = \frac{R}{RLCs^2 + Ls + R} \quad (22)$$

For a **real physical** system the function $H(s)$ is a rational polynomial such that

$$H(s) = \frac{N(s)}{D(s)}$$

with the order of the denominator $D(s)$ being greater than or equal to the order of the numerator polynomial $N(s)$. Such a system is called **proper**. In a proper system, the **system order** is the degree of the denominator polynomial.

If the denominator polynomial is extracted and set equal to zero, that is: $D(s) = 0$, the resulting equation is called **the system's characteristic equation**, since it can be shown to characterize the system's dynamics. It's roots are the **poles** of $H(s)$. The roots of the numerator polynomial are the **zeros** of the transfer function. Both the poles and zeros of H(s) can be complex values, $s = \sigma + j\omega$.

For many design applications it is desirable to present the transfer function in a factored form:

$$H(s) = \frac{k(s - z_1)(s - z_2)...(s - z_m)}{s^r(s - p_1)(s - p_2)...(s - p_n)}$$

where $m \leq n$, p_i and z_i are the poles and zeros of the transfer function, r - the number of poles at the origin, $n + r$ - the order of the system.

Other forms

$$H(s) = \frac{k}{s^r} \frac{\prod_{j=1}^{m_1}(T_j s + 1) \prod_{j=1}^{m_2}(\frac{1}{\omega_{nj}^2}s^2 + \frac{2\zeta_j}{\omega_{nj}}s + 1)}{\prod_{j=1}^{n_1}(T_j s + 1) \prod_{j=1}^{n_2}(\frac{1}{\omega_{nj}^2}s^2 + \frac{2\zeta_j}{\omega_{nj}}s + 1)}$$

where k - the gain factor, ω_{nj} - the natural frequencies, T_j - the time constants, ζ_j - the damping factors.

7 System response

The transfer function can be easily used for the calculation of system response to an input signal. Given the input signal $r(t)$ and the transfer function model of a system $H(s)$, we are interested in finding the system behavior, that is the output signal $y(t)$ (see Figure 7). From the definition of the transfer function:

$$Y(s) = H(s) \cdot R(s) \quad (23)$$

By applying the inverse Laplace transform we obtain:

$$y(t) = \pounds^{-1}[H(s) \cdot R(s)]. \quad (24)$$

Example. Spring-mass-damper response. Consider the spring-mass-damper system from the previous example and Figure 2. The transfer function of this system is given by:

$$H(s) = \frac{Y(s)}{R(s)} = \frac{1}{Ms^2 + fs + k}.$$

We shall obtain the system response or output variable, when the input r(t) is a Dirac impulse and the system constants M, f, k are:

1. $M = 1, f = 3, k = 2$
2. $M = 1, f = 1, k = 3$

If $r(t)$ is a unit impulse, the Laplace transform will be $R(s) = 1$ and the output variable $y(t)$ is:

$$y(t) = \pounds^{-1}[H(s)]$$

1. In the first case the transfer function has two real poles

   $$Y(s) = \frac{1}{(s+1)(s+2)}$$

   Expanding $Y(s)$ in a partial fraction expansion, we obtain

   $$Y(s) = \frac{k_1}{s+1} + \frac{k_2}{s+2}$$

   where k_1 and k_2 are the coefficients of the expansion, or the *residues*, and are evaluated by multiplying through by the denominator factor corresponding to k_i and setting s equal to the root.

   $$k_1 = \frac{1}{s+2}|_{s=-1} = 1$$
   $$k_2 = \frac{1}{s+1}|_{s=-2} = -1$$

   The inverse Laplace transform of $Y(s)$ is then

   $$y(t) = \pounds^{-1}\left[\frac{1}{s+1}\right] - \pounds^{-1}\left[\frac{1}{s+2}\right]$$

   Using a Laplace transforms table, we find that

   $$y(t) = e^{-t} - e^{-2t}$$

   The response $y(t)$ is called overdamped and is shown in Figure 9.

2. In this case $H(s)$ has complex poles and can be written as:

   $$H(s) = \frac{K}{\frac{1}{\omega_n^2}s^2 + \frac{2\zeta}{\omega_n}s + 1}$$

   where $K = 1/k$, $1/\omega_n^2 = M/k$ and $2\zeta/\omega_n = f/k$. It can be easily proved that when $\zeta > 1$ the roots are real, when $\zeta < 1$ the roots are complex and conjugates. When $\zeta = 1$ the roots are real and repeated and the condition is called critical damping. In this case, the damping factor $\zeta = 1/2\sqrt(3) < 1$ so the system response will be called underdamped as the poles of the transfer function are complex.

$$y(t) = \pounds^{-1}[H(s)]$$

Using a table with Laplace transform properties, and writing $Y(s)$ as

$$Y(s) = \frac{2}{\sqrt{11}} \frac{\frac{\sqrt{11}}{2}}{(s+\frac{1}{2})^2 + (\frac{\sqrt{11}}{2})^2}$$

the system response is:

$$y(t) = \frac{2}{\sqrt{11}} e^{-t/2} sin(\frac{\sqrt{11}}{2} t)$$

The underdamped response is shown in Figure 9.

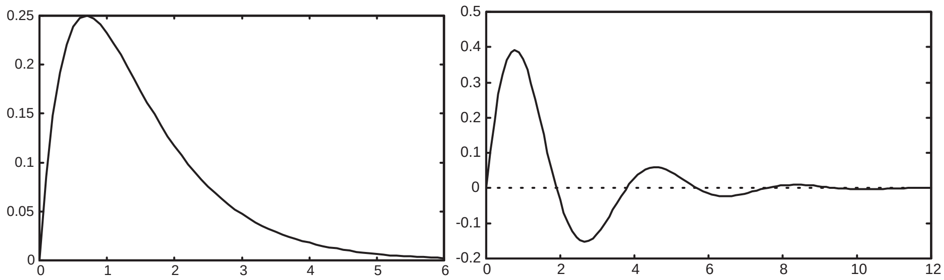

Figure 9: Spring-mass-damper response. Overdamped case (left). Underdamped case (right)

8 Block Diagram Models

Since control systems are concerned with the control of specific variables, the interrelationship of the control variables to the controlling variables is required. This relationship is typically represented by the transfer function of the subsystem relating the input and output variables. Therefore one can correctly assume that the transfer function is an important relation for control engineering.

The importance of the cause-effect relationship of the transfer function is evidenced by the facility to represent the relationship of the system variables by schematic means. The **block diagram** representation of the system relationship is prevalent in control system engineering. Block diagrams consist of *unidirectional*, operational blocks that represent the transfer functions of the variables of interest, [1].

The block diagram representation of a given system can often be reduced, using block diagram reduction techniques. The resulting diagram obeys the law of superposition because of the linear behavior of each subsystem.

In control engineering, the block diagram is a primary tool that together with transfer functions can be used to describe cause-and-effect relationships throughout a dynamic system. The manipulation of block diagrams adheres to a mathematical system of rules often known as block diagram algebra [2].

In general, the block diagram of a linear time invariant system consists of four components, namely signal, block (with transfer function), summing point and pickoff point as shown in Figure 10, [2].

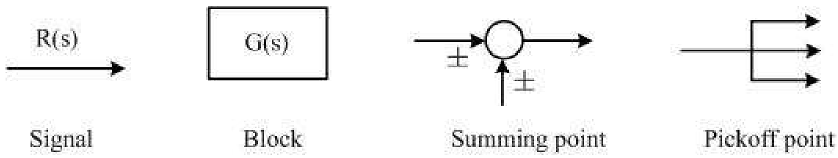

Figure 10: The components of a block diagram

In Figure 11 are shown the three basic connections: series (cascade), parallel and feedback. The equivalent transfer function for each of these cases can be determined as follows:

1. Series connection.

$$H_1(s) = \frac{Y_1(s)}{R_1(s)};\ H_2(s) = \frac{Y_2(s)}{R_2(s)};\ Y_1(s) = R_2(s)$$

The equivalent transfer function can be calculated from:

$$H(s) = \frac{Y(s)}{R(s)} = \frac{Y_2(s)}{R_1(s)} = \frac{Y_2(s) \cdot Y_1(s)}{R_1(s) \cdot R_2(s)} = H_1(s) \cdot H_2(s)$$

For n systems connected in cascade the equivalent transfer function will be:

$$H(s) = \prod_{j=1}^{n} H_j(s)$$

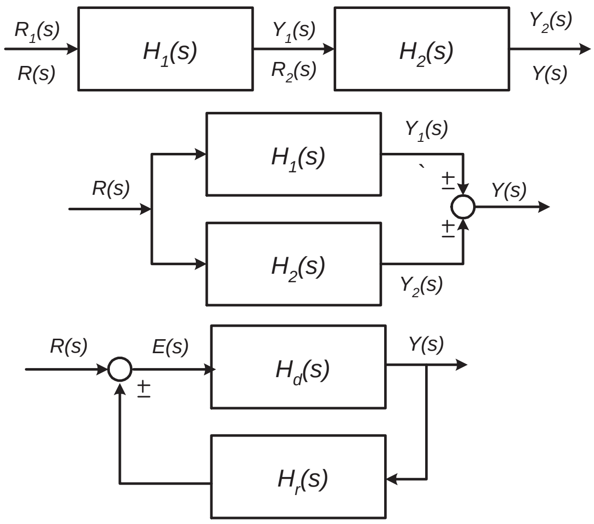

Figure 11: Three basic connections. 1-series, 2-parallel, 3-feedback

2. Parallel connection.

$$H_1(s) = \frac{Y_1(s)}{R(s)};\ H_2(s) = \frac{Y_2(s)}{R(s)};\ Y(s) = \pm Y_1(s) \pm Y_2(s)$$

The equivalent transfer function is:

$$H(s) = \frac{Y(s)}{R(s)} = \pm H_1(s) \pm H_2(s)$$

For n systems connected in parallel the equivalent transfer function is:

$$H(s) = \sum_{j=1}^{n} H_j(s)$$

3. Feedback connection. The error signal $E(s)$ can be calculated from:

$$E(s) = R(s) \pm H_r(s) \cdot Y(s)$$

The output is:

$$Y(s) = H_d(s) \cdot E(s)$$

If $E(s)$ is eliminated between the previous two relations, the overall transfer function will be:

$$H(s) = \frac{Y(s)}{R(s)} = \frac{H_d(s)}{1 \mp H_d(s) \cdot H_r(s)}$$

Diagram transformations

1. Point behind a block - Figure 12
2. Moving a pickoff point ahead of a block - Figure 13
3. Moving a pickoff point behind a block - Figure 14
4. Moving a summing point ahead of a block - Figure 15

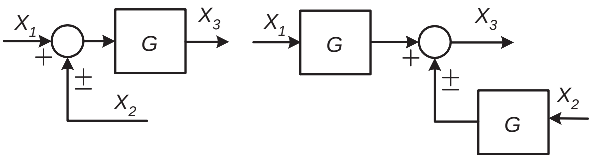

Figure 12: Point behind a block

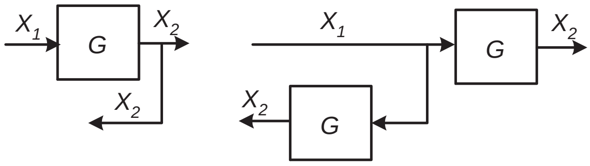

Figure 13: Moving a pickoff point ahead of a block

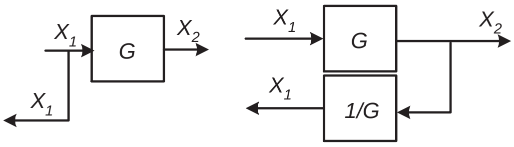

Figure 14: Moving a pickoff point behind a block

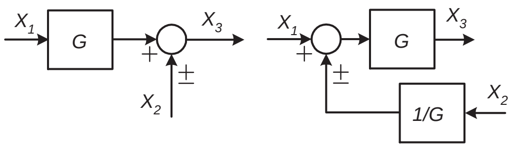

Figure 15: Moving a summing point ahead of a block

8.1 The overlap of signals

If a system has more than one input, than the output can be considered a result of all effects of these signals.

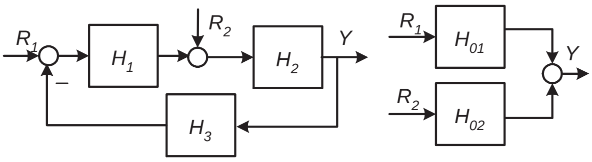

Figure 16: Two-input system

Consider the system shown in Figure 16. A transfer function relating each input and the output can be calculated, when the other input is considered zero.

$$Y(s) = R_1(s) \cdot H_{01}(s)|_{R_2(s)=0} + R_2(s) \cdot H_{02}(s)|_{R_1(s)=0}$$

$$Y(s) = \frac{H_1 H_2}{1 + H_1 H_2 H_3} \cdot R_1(s) + \frac{H_2}{1 + H_1 H_2 H_3} \cdot R_2(s)$$

8.2 Example

Consider a system represented by the block diagram shown in Figure 17. Determine the closed-loop transfer function $T(s) = Y(s)/R(s)$.

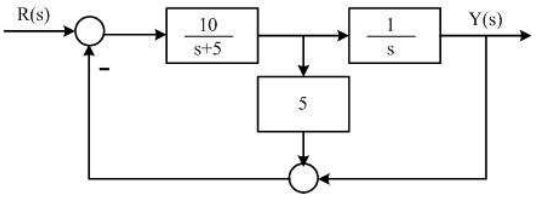

Figure 17: Block diagram example exercise.

The block diagram will be transformed so the basic rules of block diagram reduction can be applied. First, the summing points encircled in Figure 18 (left) will be considered. Note that the signal Z can be calculated as:

$$Z = R - (X + Y) = R - X - Y \tag{25}$$

and the two summing points can be reduced to a single one, as shown in Figure 18 (right)

From now on, the basic rules apply. The inner loop is reduced first to one block and the equivalent transfer function is:

$$G_{in}(s) = \frac{\frac{10}{s+5}}{1 + \frac{10}{s+5} \cdot 5} = \frac{10}{s + 55} \tag{26}$$

This resulting block, is in a series connection with $1/s$ and the equivalent of these two is the product G_{in}/s. The new block diagram is presented in Figure 19.

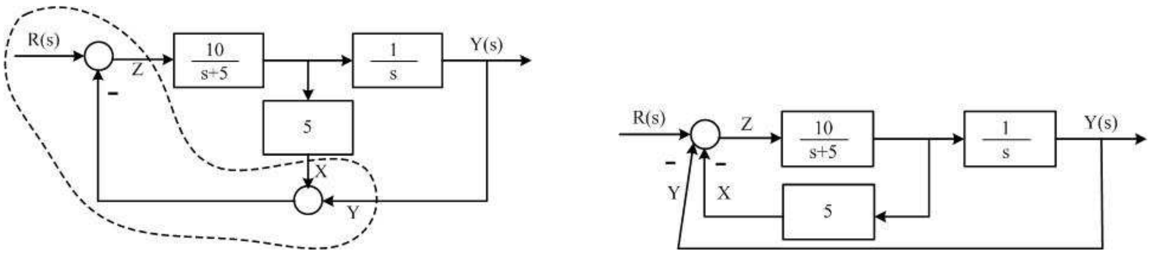

Figure 18: Block diagram transformation.

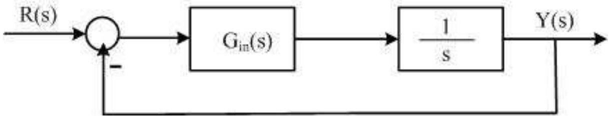

Figure 19: Block diagram transformation.

Note that the feedback path has a transfer function equal to 1, therefore it is a *unity feedback* loop. The equivalent transfer function from the input $R(s)$ to the output $Y(s)$ is calculated as follows:

$$T(s) = \frac{Y(s)}{R(s)} = \frac{G_{in}(s)\cdot\frac{1}{s}}{1+G_{in}(s)\cdot\frac{1}{s}\cdot 1} = \frac{\frac{10}{s(s+55)}}{1+\frac{10}{s(s+55)}} = \frac{10}{s^2+55s+10} \tag{27}$$

9 The Transfer Matrix

Because of the great number of systems with multiple inputs and outputs (MIMO systems), calculus techniques specific for this kind of systems were developed. In case of linear systems the mathematical techniques use matrix calculus instead of transfer functions.

In Figure 20 the block diagram of a system with m inputs and n outputs is shown.

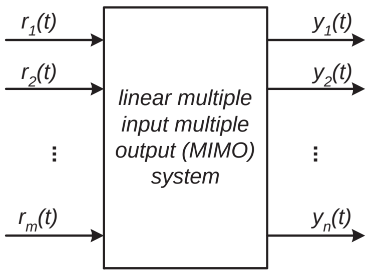

Figure 20: MIMO system

If the signal transfer is unidirectional and each input influences each output, as shown in Figure 21, we can write the following equation system:

$$\begin{gathered} Y_1 = H_{11}R_1 + H_{12}R_2 + \ldots H_{1m}R_m \\ Y_2 = H_{21}R_1 + H_{22}R_2 + \ldots H_{2m}R_m \\ \ldots \\ Y_n = H_{n1}R_1 + H_{n2}R_2 + \ldots H_{nm}R_m \end{gathered}$$

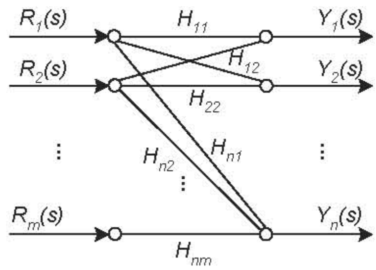

Figure 21: Transfer functions between inputs and outputs

where R and Y are the Laplace transforms of the inputs $r(t)$ and outputs $y(t)$, and

$$H_{jk} = \frac{Y_j}{R_k}$$

is the transfer function from the input k to the output j.

Thus, a multiple linear system is composed by nxm elements having one input and one output. The equation system can be written in a matrix form:

$$\mathbf{Y} = \mathbf{H} \cdot \mathbf{R}$$

where

$$\mathbf{R} = [R_1(s)\; R_2(s)\; ...\; R_m(s)]^T$$

is the input vector (m x 1),

$$\mathbf{Y} = [Y_1(s)\; Y_2(s)\; ...\; Y_n(s)]^T$$

is the output vector (n x 1), and

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} & \dots & H_{1m} \\ H_{21} & H_{22} & \dots & H_{2m} \\ \dots & \dots & \dots & \dots \\ H_{n1} & H_{n2} & \dots & H_{nm} \end{bmatrix}$$

is called the **transfer matrix**.

In case of connecting linear MIMO systems, we can use transfer matrices instead of transfer functions, obeying the rules of matrix calculus.

For example consider the three basic connections shown in Figure 22.

1. Series connection.

   $$\mathbf{Y}_1 = \mathbf{H}_1 \cdot \mathbf{R};\;\; \mathbf{Y} = \mathbf{H}_2 \cdot \mathbf{R}_2;\;\; \mathbf{R}_2 = \mathbf{Y}_1;\;\; \Rightarrow\;\; \mathbf{Y} = \mathbf{H}_2 \cdot \mathbf{H}_1 \cdot \mathbf{R}$$

   The equivalent transfer matrix will be:

   $$\mathbf{H} = \mathbf{H}_2 \cdot \mathbf{H}_1$$

   For n linear systems connected in series, the equivalent transfer matrix is:

   $$\mathbf{H} = \prod_{j=n}^{1} \mathbf{H}_j$$

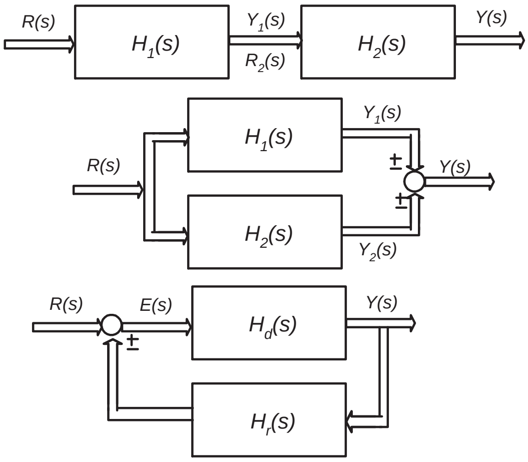

Figure 22: MIMO system

2. Parallel connection.

$$\mathbf{Y}_1 = \mathbf{H}_1 \cdot \mathbf{R}; \quad \mathbf{Y}_2 = \mathbf{H}_2 \cdot \mathbf{R}; \quad \mathbf{Y} = \pm\mathbf{Y}_1 \pm \mathbf{Y}_2$$

The equivalent transfer matrix will be:

$$\mathbf{H} = \pm\mathbf{H}_1 \pm \mathbf{H}_2$$

3. Feedback connection.

$$\mathbf{E} = \mathbf{R} \pm \mathbf{H}_r \cdot \mathbf{Y}; \quad \mathbf{Y} = \mathbf{H}_d \cdot \mathbf{E}; \quad \mathbf{Y} = \mathbf{H}_d \cdot (\mathbf{R} \pm \mathbf{H}_r \cdot \mathbf{Y})$$

$$(\mathbf{1} \mp \mathbf{H}_d \cdot \mathbf{H}_r) \cdot \mathbf{Y} = \mathbf{H}_d \cdot \mathbf{R}; \quad \mathbf{Y} = (\mathbf{1} \mp \mathbf{H}_d \cdot \mathbf{H}_r)^{-1} \cdot \mathbf{H}_d \cdot \mathbf{R}$$

The equivalent transfer matrix will be:

$$\mathbf{H} = (\mathbf{1} \mp \mathbf{H}_d \cdot \mathbf{H}_r)^{-1} \cdot \mathbf{H}_d$$

References

[1] Richard C. Dorf and Robert H. Bishop, *Modern Control Systems*, Addison-Wesley, 1995.

[2] C. Mei, *On teaching the simplification of block diagrams*, International Journal of Engineering Education **18** (2002), no. 6, 697–703.