Chapter 2

State-space and transfer functions

TODO: include examples and simulations

2.1 State-space transformations

Recall that the standard state-space description of a linear continuous system is given by a **state equation** that relates the rate of change of the state of the system to the state of the system and the input signals and the **output equation** where the outputs are related to the state variables and the input signals. The state-space model is a set of two matrix equations written in the compact form as:

$$\begin{cases} \dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} = C\mathbf{x} + D\mathbf{u} \end{cases} \tag{2.1}$$

where the state vector $\mathbf{x}$, the input vector **u** and the output vector **y** are:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n_x} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n_u} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n_y} \end{bmatrix} \tag{2.2}$$

In equations (2.1):

- $\mathbf{x}$ is an $(n_x \times 1)$ state vector, where n_x is the number of states or system order
- **u** is an $(n_u \times 1)$ input vector, where n_u is the number of input functions
- $\mathbf{y}$ is a $(n_y \times 1)$ output vector where n_y is the number of outputs
- A is an $(n_x \times n_x)$ square matrix called the **system matrix** or **state matrix**
- B is an $(n_x \times n_u)$ matrix called the **input matrix**

- C is a $(n_y \times n_x)$ matrix called the **output matrix**
- D is a $(n_y \times n_u)$ matrix which represents any direct connection between the input and the output. It as called the **feedthrough** matrix.

It has been already shown in the previous chapter that the definition of a state in a state-space description is not unique. This can be seen by considering the linear time-invariant system

$$\begin{aligned} \dot{\boldsymbol{x}} &= A\boldsymbol{x} + B\boldsymbol{u} \\ \boldsymbol{y} &= C\boldsymbol{x} + D\boldsymbol{u} \end{aligned} \tag{2.3}$$

where $\boldsymbol{x}$ denotes the state vector, $\boldsymbol{u}$ is the input vector, $\boldsymbol{y}$ is the output (measurement vector), A is the state matrix, B is the input matrix, C is the output and D is the feedthrough matrix, and a linear transformation of $\boldsymbol{x}$ to $\bar{\boldsymbol{x}}$ defined as

$$\bar{\boldsymbol{x}} = T^{-1}\boldsymbol{x} \qquad \boldsymbol{x} = T\bar{\boldsymbol{x}}$$

where T is any nonsingular matrix of appropriate dimensions. Substituting the new variables into (2.3) we obtain

$$\begin{aligned} \dot{\bar{\boldsymbol{x}}} &= \bar{A}\bar{\boldsymbol{x}} + \bar{B}\boldsymbol{u} \\ \boldsymbol{y} &= \bar{C}\bar{\boldsymbol{x}} + \bar{D}\boldsymbol{u} \end{aligned} \tag{2.4}$$

with $\bar{A} = T^{-1}AT$, $\bar{B} = T^{-1}B$, $\bar{C} = CT$, $\bar{D} = D$.

Note that only the state vector and the matrices have changed, the input and output remain the same. Since the transfer function does not depend on the state vector, the same transfer function is obtained, given by

$$H(s) = \frac{Y(s)}{U(s)} \tag{2.5}$$

$H(s)$ denotes the transfer function in the s-domain, and $Y(s)$ and $U(s)$ are the Laplace transform of $y(t)$ and $u(t)$, assuming zero initial conditions.

2.2 From state-space to transfer matrix

To see how the transfer function is obtained, consider the Laplace transform of (2.3):

$$\begin{aligned} sX(s) - X(0) &= AX(s) + BU(s) \\ Y(s) &= CX(s) + DU(s) \end{aligned}$$

Reordering the terms we get:

$$\begin{aligned} (sI - A)X(s) &= BU(s) + X(0) \\ X(s) &= (sI - A)^{-1}(X(0) + BU(s)) \end{aligned}$$

and the output can be written as

$$Y(s) = C(sI - A)^{-1}(X(0) + BU(s)) + DU(s)$$

or

$$Y(s) = C\left((sI - A)^{-1}B + D\right)U(s) + C(sI - A)^{-1}X(0)$$

which, as can be seen, depends on the initial conditions. In zero initial conditions, we obtain the well-known expression

$$Y(s) = \underbrace{C\left((sI - A)^{-1}B + D\right)}_{H(s):=}U(s)$$

and the transfer function

$$H(s) = C(sI - A)^{-1}B + D$$

Now, if the transfer function is computed using the transformed matrices $\bar{A}$, $\bar{B}$, $\bar{C}$, and $\bar{D}$, we get

$$\begin{aligned}\bar{H}(s) &= CT(sI - T^{-1}AT)^{-1}T^{-1}B + D\\ &= CTT^{-1}(sI - A)^{-1}TT^{-1}B + D\\ &= C(sI - A)^{-1}B + D\end{aligned}$$

i.e., the same transfer function has been obtained as before.

Note that since

$$H(s) = \frac{C\,\mathrm{Adj}(sI - A)B}{\det(sI - A)} + D$$

the system poles are given by solution of the characteristic equation $\det(sI - A) = 0$, are also the eigenvalues of A.

2.3 From transfer function to state-space

This section is taken from Chapter 2 of **?**.

Although the transformation from transfer function to a state-space model is not unique, here we present a method to obtain the state variables in the form of *phase variables*. The state variables are *phase variables* when each subsequent state is defined to be the derivative of the previous state variable.

Consider a system with the input $u(t)$ and the output $y(t)$ described by the n-th order linear differential equation:

$$\frac{d^n y(t)}{dt^n} + a_{n-1}\frac{d^{n-1}y(t)}{dt^{n-1}} + \cdots + a_1\frac{dy(t)}{dt} + a_0 y(t) = b_0 u(t) \qquad (2.6)$$

A convenient way to chose the state variables is to choose the output $y(t)$ and its $n-1$ derivatives as the state variables. They are called *phase variables*:

$$\begin{aligned} x_1 &= y \\ x_2 &= \frac{dy}{dt} \\ &\vdots \\ x_n &= \frac{d^{n-1}y}{dt^{n-1}} \end{aligned} \tag{2.7}$$

Differentiating both sides of the system (2.7) yields:

$$\dot{x}_n = \frac{d^n y}{dt^n} \tag{2.8}$$

If we denote $\dot{x}_i = \frac{d^i x}{dt^i}$ the system (2.7) can be written also as:

$$\begin{aligned} x_1 &= y \\ x_2 &= \frac{dy}{dt} = \frac{dx_1}{dt} = \dot{x}_1 \\ x_3 &= \frac{d^2y}{dt^2} = \frac{dx_2}{dt} = \dot{x}_2 \\ &\vdots \\ x_n &= \frac{d^{n-1}y}{dt^{n-1}} = \frac{dx_{n-1}}{dt} = \dot{x}_{n-1} \end{aligned} \tag{2.9}$$

Substituting the definitions (2.7) and (2.8) into (2.6) we obtain:

$$\dot{x}_n + a_{n-1}x_n + \cdots + a_1 x_2 + a_0 x_1 = b_0 u \tag{2.10}$$

The n-th order differential equation (2.6) is equivalent to a system of n first order differential equations obtained from the definitions of the derivatives from (2.9) together with the $\dot{x}_n$ that results from (2.10):

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= -a_0 x_1 - a_1 x_2 - \cdots - a_{n-1}x_n + b_0 u \end{aligned}$$

In a matrix-vector form equations (2.11) become:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 & \cdots & -a_{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ b_0 \end{bmatrix} u \tag{2.11}$$

Equation (2.11) is the *phase-variable* form of the state equation. This form is easily recognized by the pattern of 1's above the main diagonal and 0's for the rest of the state matrix, except for the last row that contains the coefficients of the differential equation written in reverse order, (Nise, 2004).

The output equation is:

$$y = x_1$$

or, in a vector form:

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + 0 \cdot u \tag{2.12}$$

2.3.1 Converting a transfer function with constant term at numerator

For a system with an input u and an output y consider a general transfer function with constant term at numerator:

$$H(s) = \frac{b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0} = \frac{Y(s)}{U(s)}$$

Cross-multiplying the relation above yields:

$$(s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0)Y(s) = b_0 U(s)$$

and by taking the inverse Laplace transform we get:

$$\frac{d^n y}{dt^n} + a_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \cdots + a_1 \frac{dy}{dt} + a_0 y = b_0 u$$

This is exactly the equation (2.6) for which the phase-variable form was obtained above. The state equation is then (2.11) and the output equation is (2.12). Note that the matrix $\mathbf{D} = 0$.

Example 2.1 *Consider a system with the input u and an output y described by the transfer function:*

$$H(s) = \frac{2}{s^3 + 2s^2 + 3s + 4}$$

The transfer function written as the ratio of the Laplace transform of the output to the Laplace transform of the input, with all the initial conditions assumed to be zero is:

$$\frac{Y(s)}{U(s)} = \frac{2}{s^3 + 2s^2 + 3s + 4}$$

or

$$(s^3 + 2s^2 + 3s + 4)Y(s) = 2U(s)$$

Taking the inverse Laplace transform we obtain the differential equation:

$$\frac{d^3y}{dt^3} + 2\frac{d^2y}{dt^2} + 3\frac{dy}{dt} + 4y = 2u \tag{2.13}$$

Choosing the state variables as successive derivatives (phase variable form) we get:

$$\begin{aligned}
x_1 &= y \\
x_2 &= \frac{dy}{dt} = \frac{dx_1}{dt} = \dot{x}_1 \\
x_3 &= \frac{d^2y}{dt^2} = \frac{dx_2}{dt} = \dot{x}_2
\end{aligned}$$

and the derivative of the last state variable is:

$$\dot{x}_3 = \frac{d^3y}{dt^3}$$

All the above definitions are now replaced into (2.13):

$$\dot{x}_3 + 2x_3 + 3x_2 + 4x_1 = 2u$$

This last relation and the definitions of the phase variables will give the state equation in the phase variable form:

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_3 \\
\dot{x}_3 &= -4x_1 - 3x_2 - 2x_3 + 2u
\end{aligned}$$

and the output equation:

$$y = x_1$$

In the matrix-vector form, the state-space model is:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & -3 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} u \tag{2.14}$$

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{2.15}$$

If the vector of state variables is denoted by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

the state-space model is:

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & -3 & -2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{x}$$

2.3.2 Converting a transfer function with polynomial at numerator

Although the method presented below can be applied for systems of any order, to simplify the demonstration, consider a third-order transfer function with a second-order polynomial in the numerator:

$$H(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0}$$

Notice that *the denominator is a monic polynomial* (the leading coefficient or the coefficient of highest degree is equal to 1). If the polynomial in the numerator is of order less than the polynomial in the denominator, the numerator and denominator can be handled separately. First, separate the transfer function into two cascaded transfer functions, as shown in Figure 2.1: the first is the denominator and the second one is just the numerator, (Nise, 2004).

The first transfer function will be converted to the phase-variable representation in state-space as demonstrated in the previous subsection 2.3.1. Hence, phase variable x_1 is the output and the rest of the phase variables are the internal variables of the first block as shown in Figure 2.1.

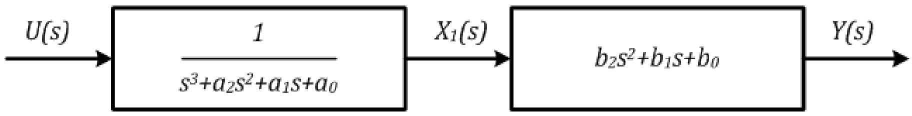

Figure 2.1: Decomposing a transfer function

The first transfer function is:

$$\frac{X_1(s)}{U(s)} = \frac{1}{s^3 + a_2 s^2 + a_1 s + a_0} \tag{2.16}$$

and the second one:

$$\frac{Y(s)}{X_1(s)} = b_2 s^2 + b_1 s + b_0 \tag{2.17}$$

Following the procedure described in the previous section, the state equation resulting from (2.16) will be:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u \tag{2.18}$$

The second transfer function with just the numerator yields:

$$Y(s) = (b_2 s^2 + b_1 s + b_0) X_1(s)$$

where, after taking the inverse Laplace transform with zero initial conditions:

$$y = b_2 \frac{d^2 x_1}{dt^2} + b_1 \frac{dx_1}{dt} + b_0 x_1$$

But the derivative terms are the definitions of the phase variables obtained in the first block. Thus, writing the terms in reverse order, the output equation is:

$$y = b_0 x_1 + b_1 x_2 + b_2 x_3$$

or, in a matrix form:

$$y = \begin{bmatrix} b_0 & b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{2.19}$$

Hence, the second block simply forms a specified linear combination of the state variables developed in the first block.

From another perspective, the denominator of the transfer function yields the state equations while the numerator yields the output equation.

If the order of the polynomial in the numerator is equal to the order of the polynomial in the denominator, the third-order transfer function will be written in the general form:

$$H(s) = \frac{b_3 s^3 + b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0}$$

and the numerator resulted from the decomposition is:

$$\frac{Y(s)}{X_1(s)} = b_3 s^3 + b_2 s^2 + b_1 s + b_0 \text{ or } Y(s) = (b_3 s^3 + b_2 s^2 + b_1 s + b_0) X_1(s)$$

After taking the inverse Laplace transform with zero initial conditions:

$$y = b_3 \frac{d^3 x_1}{dt^3} + b_2 \frac{d^2 x_1}{dt^2} + b_1 \frac{dx_1}{dt} + b_0 x_1$$

Using the definitions of the state variables resulted in the first block and re-arranging the terms:

$$y = b_0 x_1 + b_1 x_2 + b_2 x_3 + b_3 \dot{x}_3$$

The output equation must not contain any derivatives of the state variables, but we can replace $\dot{x}_3$ with the last equation of the system (2.18):

$$\dot{x}_3 = -a_0 x_1 - a_1 x_2 - a_2 x_3 + u$$

so it will become:

$$y = b_0 x_1 + b_1 x_2 + b_2 x_3 + b_3(-a_0 x_1 - a_1 x_2 - a_2 x_3 + u)$$

Re-arranging the output equation is:

$$y = (b_0 - b_3 a_0) x_1 + (b_1 - b_3 a_1) x_2 + (b_2 - b_3 a_2) x_3 + b_3 u$$

or, in the matrix form:

$$y = \begin{bmatrix} b_0 - b_3 a_0 & b_1 - b_3 a_1 & b_2 - b_3 a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + [b_3]\, u \qquad (2.20)$$

Note that the output equation (2.19) is the same as (2.20) for $b_3 = 0$. Also, note that the matrix $\mathbf{D}$ is no longer equal to zero. If the system has one input and one output, D is a scalar value and in this case $D = b_3$.

If the system is represented by a transfer function, the minimum number of state variables that have to be chosen is equal to the order of the system (or the order of the polynomial in the denominator of the transfer function).

Example 2.2 *Consider a system with the input u and an output y described by the transfer function:*

$$H(s) = \frac{s+2}{s^2+2s+2}$$

Determine the state-space model in the phase variable form.

- *Separate the transfer function into two cascaded blocks.*

  The transfer functions written as the ratio of the Laplace transform of the output to the Laplace transform of the input, with all the initial conditions assumed to be zero are:

  $$\frac{X_1(s)}{U(s)} = \frac{1}{s^2+2s+2}, \quad \frac{Y(s)}{X_1(s)} = s+2$$

- *Find the state equations for the first block. From the first transfer function we have:*

  $$(s^2+2s+2)X_1(s) = U(s)$$

  and the differential equation obtained by taking the inverse Laplace transform is:

  $$\ddot{x}_1 + 2\dot{x}_1 + 2x_1 = u$$

  The first state variable was already chosen as the output of the first block x_1 and the number of state variables is 2, equal to the order of the system. Therefore we choose the second state in the phase variable form:

  $$x_2 = \dot{x}_1$$

  Replacing in the differential equation we obtain:

  $$\dot{x}_2 + 2x_2 + 2x_1 = u$$

  Re-arranging and taking also the definition of the second state variable we obtain the state equations:

  $$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -2x_1 - 2x_2 + u \end{aligned}$$

  or

  $$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

  or

  $$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \tag{2.21}$$

- *Introduce now the block with the numerator. From the transfer function we obtain:*

  $$Y(s) = (s+2)X_1(s)$$

  or, in time-domain:

  $$y = \dot{x}_1 + 2x_1 = x_2 + 2x_1$$

  The output equation can be written also as:

  $$y = \left[\begin{array}{cc} 2 & 1 \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right]$$

  or

  $$y = \left[\begin{array}{cc} 2 & 1 \end{array}\right] \mathbf{x} \tag{2.22}$$