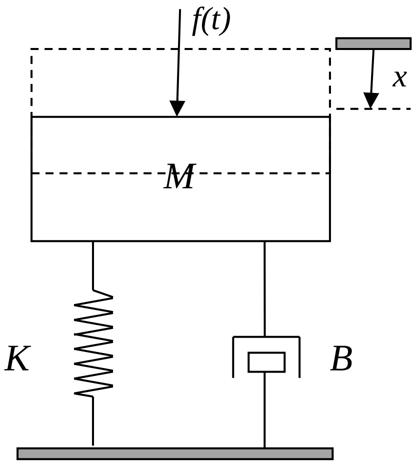
4. State-space representation

4.1 Introduction

- When the differential equations governing the system are linear, the model can be described in terms of a Laplace transfer function.

- Consider for example the spring-mass damper system:



Physical model

$$M\frac{d^2x}{dt^2} + B\frac{dx}{dt} + Kx = f(t)$$

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{Ms^2 + Bs + K}$$

- While transfer functions are compact and have several advantages in relation to analysis of dynamical systems, they also have a number of weaknesses as follows:
    - They do not handle initial conditions (assumed to be zero).
    - Information about internal variables is lost ($\frac{dx}{dt}$ in the case of the above example).
    - For general *m*-input, *p*-output systems, we would need a total of *m* x *p* transfer functions to fully describe the system.

- An alternative model representation, known as a **state-space model**, overcomes these weaknesses.

- The basic idea behind state-space modelling is to write down a set of first-order differential equations in terms of the system state(s) and input(s).

- **Example 4.1:** Develop a state-space model for the second-order differential equation model of the spring-mass damper system:

$$M\frac{d^2x}{dt^2}+B\frac{dx}{dt}+Kx=f(t)$$

Solution
We proceed as follows:

Firstly we have a ***second-order*** differential equation; hence we define two states x_1 and x_2:

$$x_1=x,\quad x_2=\frac{dx}{dt}$$

When we do it this way we get the CONTROLLABLE CANONICAL FORM state-space model.

We then obtain expressions for $\frac{dx_1}{dt}$ and $\frac{dx_2}{dt}$:

$$\frac{dx_1}{dt}=\frac{dx}{dt}\qquad\qquad \frac{dx_2}{dt}=\frac{d^2x}{dt^2}=\frac{1}{M}\left(f(t)-B\frac{dx}{dt}-Kx\right)$$

Note that the latter expression is obtained from the original mathematical model.

We can rewrite these equations in terms of the states x_1 and x_2 only, by replacing x and $\frac{dx}{dt}$ with their state equivalent. Hence:

$$\frac{dx_1}{dt}=x_2\qquad \frac{dx_2}{dt}=\frac{1}{M}\left(f(t)-Bx_2-Kx_1\right)$$

Now, we combine these into matrix form as follows:

$$\frac{d}{dt}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}0 & 1\\ -\frac{K}{M} & -\frac{B}{M}\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}+\begin{bmatrix}0\\ \frac{1}{M}\end{bmatrix}f(t)$$

Finally, let us define the output as $y = x$. Hence $y = x_1$ giving:

$$[y]=[1 \quad 0]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}+[0]f(t) \Rightarrow [y]=[1 \quad 0]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Our state-space model is expressed, in full, as:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ -\frac{K}{M} & -\frac{B}{M} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}+\begin{bmatrix} 0 \\ \frac{1}{M} \end{bmatrix}f(t)$$

$$[y]=[1 \quad 0]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

4.2 Formal definitions

- The commonly used terms associated with state-space representation are as follows:

- **State** – the state of a dynamic system is the smallest set of variables (called state variables) such that knowledge of these variables at $t = t_0$, together with knowledge of the input for $t \geq t_0$, completely determines the behaviour of the system for $t \geq t_0$.

- **State variables** – the variables that make up the state as defined above.

- **State vector** – when there is more than one state variable, they are normally collected together into a vector called a state vector.

- **State-space** – the n-dimensional state vector can be viewed as a point moving around in n-dimensional space. This n-dimensional space is known as state-space.

General state-space form

- The dynamics of a general n^{th} order linear dynamical system, with m inputs, is completely described by a n^{th} order state-space equation of the form:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ b_{21} & \cdots & b_{2m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{bmatrix}\begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

$$\equiv \quad \frac{\mathrm{d}\mathbf{x}}{\mathrm{dt}} = \mathbf{Ax} + \mathbf{Bu} \qquad \text{state equation}$$

with a set of initial conditions (one for each state):

$$\boldsymbol{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \\ \vdots \\ x_n(0) \end{bmatrix} = \boldsymbol{x_0} \qquad \text{initial conditions}$$

- The model output(s) are given by a linear combination of the states and the inputs. Given *p* outputs, we obtain:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} d_{11} & \cdots & d_{1m} \\ d_{21} & \cdots & d_{2m} \\ \vdots & \ddots & \vdots \\ d_{p1} & \cdots & d_{pm} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

$$\equiv \quad \boldsymbol{y} = \boldsymbol{Cx} + \boldsymbol{Du} \qquad \text{output equation}$$

- Together, these two equations describe the state-space model of a system.

$$\frac{d\mathbf{x}}{dt} = \mathbf{Ax} + \mathbf{Bu} \qquad , \mathbf{x}(0) = \mathbf{x_0}$$

$$\mathbf{y} = \mathbf{Cx} + \mathbf{Du}$$

- The matrices ***A***, ***B***, ***C*** and ***D*** are called the state matrix, input matrix, output matrix and direct transmission matrix respectively:

A	-	state matrix (n x n)
B	-	input matrix (n x m)
C	-	output matrix (p x n)
D	-	direct transmission matrix (p x m)

4.3 State – space representations are not unique

State-space representations are NOT unique.

Consider the following system

$$\frac{d^2y(t)}{dt^2} + 4\frac{dy(t)}{dt} + 3y(t) = 2f(t)$$

Controllable canonical form state-space representation

Form the transfer function

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

Let $x_{1c}(t) = y(t)$ and $x_{2c}(t) = \frac{dy(t)}{dt}$.

One needs two equations – one for the time derivative of state 1 and one for the time derivative of state 2.

$$\boxed{\frac{dx_{1c}}{dt} = \frac{dy(t)}{dt} = x_{2c}}$$

$$\frac{d^2y}{dt^2} + 4\frac{dy}{dt} + 3y = 2f$$

$$\frac{d^2y}{dt^2} = -4\frac{dy}{dt} - 3y + 2f$$

Now put in the states

$$\boxed{\frac{d^2y}{dt^2} = \frac{dx_{2c}}{dt} = -4x_{2c} - 3x_{1c} + 2f}$$

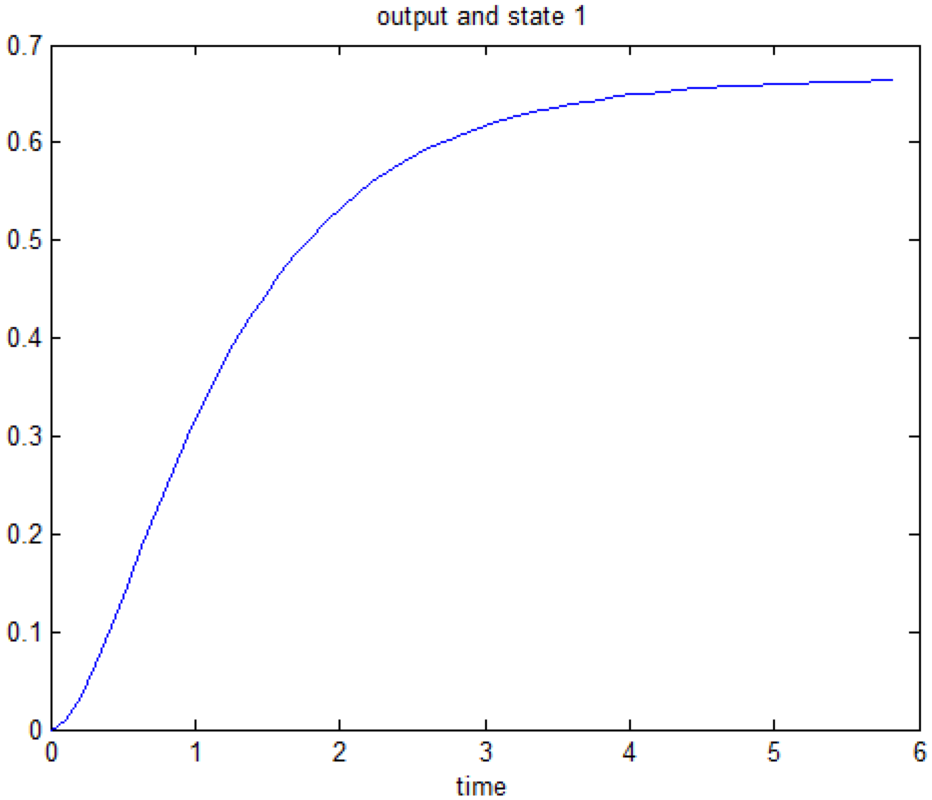
Put in matrix formation

$$\frac{d}{dt}\begin{bmatrix} x_{1c}(t) \\ x_{2c}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}\begin{bmatrix} x_{1c}(t) \\ x_{2c}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} f(t)$$

The output is $y(t)$ which is the first state

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} x_{1c}(t) \\ x_{2c}(t) \end{bmatrix}$$

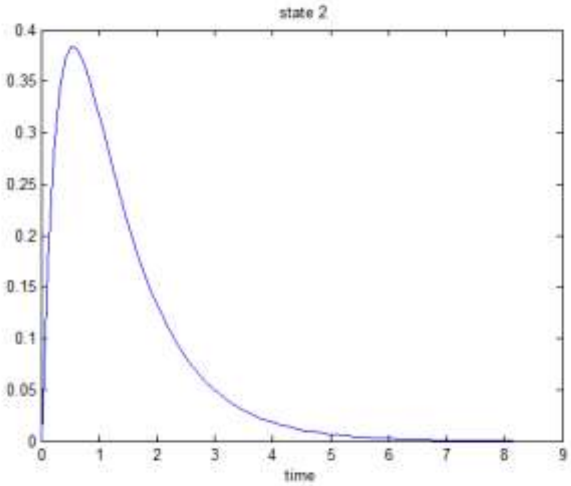
The unit-step response is as follows



Suppose you wanted to see what the second state looks like

$$o(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1c}(t) \\ x_{2c}(t) \end{bmatrix}$$

Then you could plot this.

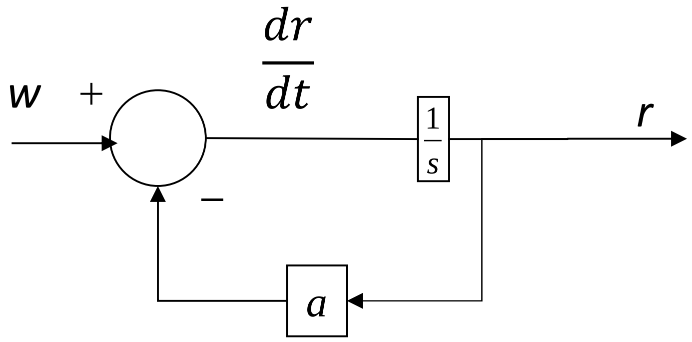


Series form representation

Express $G(s)$ as the product of $\frac{1}{s+a}$ terms.

$$G(s) = \frac{2}{s^2 + 4s + 3} = 2\left(\frac{1}{s+1}\right)\left(\frac{1}{s+3}\right)$$

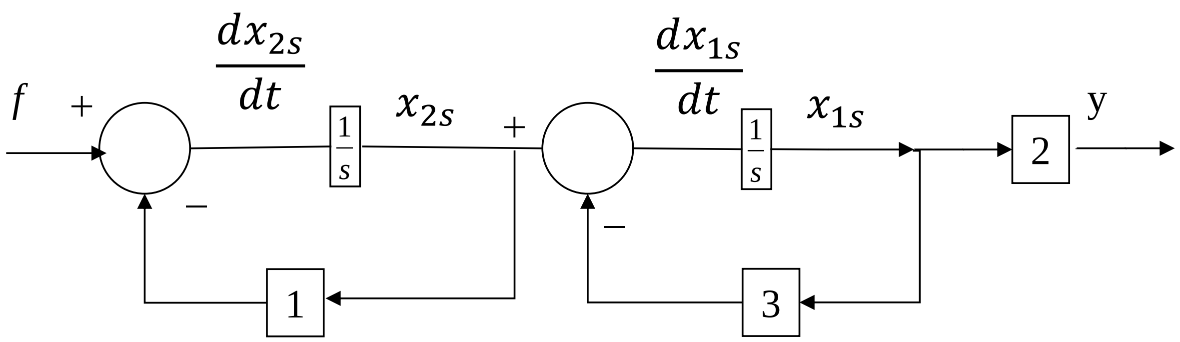
The block diagram for $\frac{1}{s+a}$ is



$$\frac{dr}{dt} = w - ar$$
$$sR(s) = W(s) - aR(s)$$
$$\frac{R(s)}{W(s)} = \frac{1}{s+a}$$

So $G(s)$ can be represented in block diagram form as follows

$$G(s) = \frac{2}{s^2 + 4s + 3} = 2\left(\frac{1}{s+1}\right)\left(\frac{1}{s+3}\right)$$



$$\frac{dx_{1s}}{dt} = -3x_{1s} + x_{2s}$$

$$\frac{dx_{2s}}{dt} = -x_{2s} + f$$
$$y = 2x_{1s}$$

Now write the state-space representation in matrix form

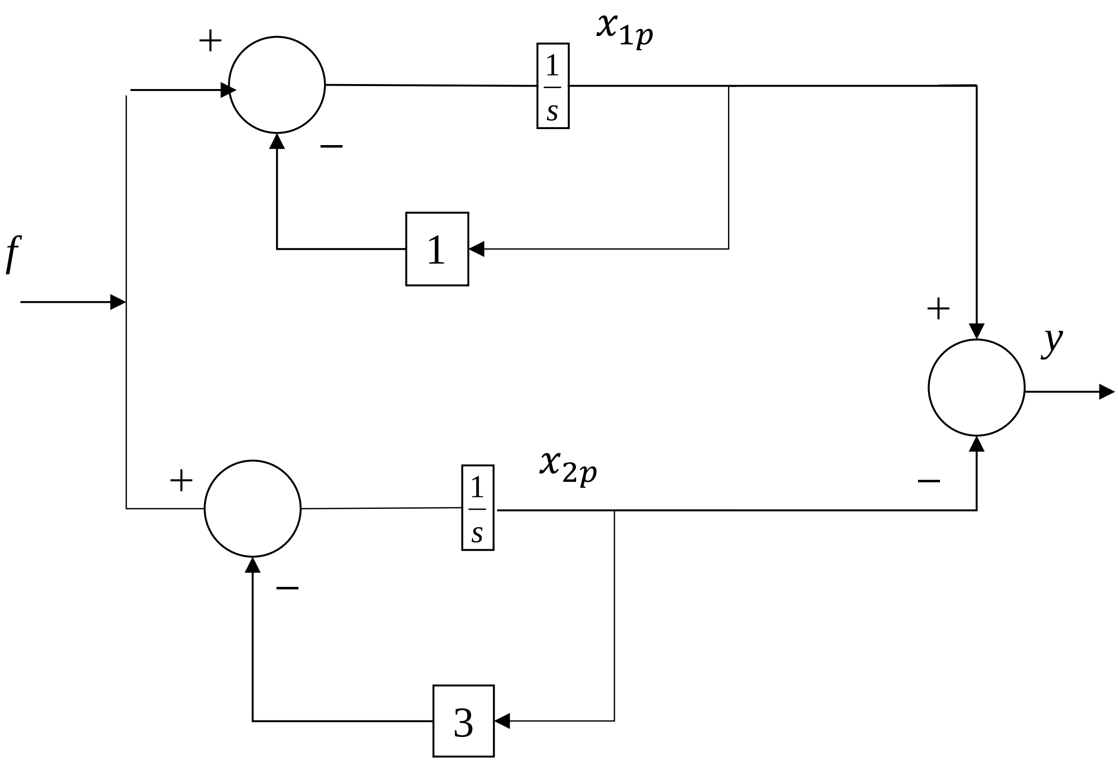
$$\frac{d}{dt}\begin{bmatrix} x_{1s}(t) \\ x_{2s}(t) \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x_{1s}(t) \\ x_{2s}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} f(t)$$
$$y(t) = \begin{bmatrix} 2 & 0 \end{bmatrix}\begin{bmatrix} x_{1s}(t) \\ x_{2s}(t) \end{bmatrix}$$

Parallel form state-space representation

Express $G(s)$ as the sum or difference of $\frac{1}{s+a}$ terms.

$$G(s) = \frac{2}{s^2 + 4s + 3} = \left(\frac{1}{s+1}\right) - \left(\frac{1}{s+3}\right)$$

This can be shown in block diagram format as follows



$$\frac{dx_{1p}}{dt} = -x_{1p} + f$$
$$\frac{dx_{2p}}{dt} = -3x_{2p} + f$$

$$\frac{d}{dt}\begin{bmatrix} x_{1p}(t) \\ x_{2p}(t) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} x_{1p}(t) \\ x_{2p}(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} f(t)$$

$$y(t) = \begin{bmatrix} 1 & -1 \end{bmatrix}\begin{bmatrix} x_{1p}(t) \\ x_{2p}(t) \end{bmatrix}$$

Note that the state matrix is diagonal. This is important for computations as solution of the state equations (to be dealt with in a later lecture) involves exponentials of the state matrix and it is much easier to get the exponential of a diagonal matrix.

4.4 Transforming system models to state-space form

Derivation of the state-space model from the transfer function model of the system.

4.4.1 Transfer function → State space

Case 1: when the input does *not* involve derivatives

- Consider the following third-order transfer function:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{k}{s^3 + a_2 s^2 + a_1 s + a_0}$$

- Converting this to the time domain gives:

$$\frac{d^3 y}{dt^3} + a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = ku$$

or

$$\frac{d^3 y}{dt^3} = -a_2 \frac{d^2 y}{dt^2} - a_1 \frac{dy}{dt} - a_0 y + ku$$

- We simply define the states as $x_1 = y, x_2 = \frac{dy}{dt}$ and $x_3 = \frac{d^2 y}{dt^2}$, leading to the following state model:

$$\frac{dx_1}{dt} = \frac{dy}{dt} = x_2$$

$$\frac{dx_2}{dt} = \frac{d^2 y}{dt^2} = x_3$$

$$\frac{dx_3}{dt} = \frac{d^3 y}{dt^3} = -a_2 x_3 - a_1 x_2 - a_0 x_1 + ku$$

$$y = x_1$$

giving:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ k \end{bmatrix} u$$

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

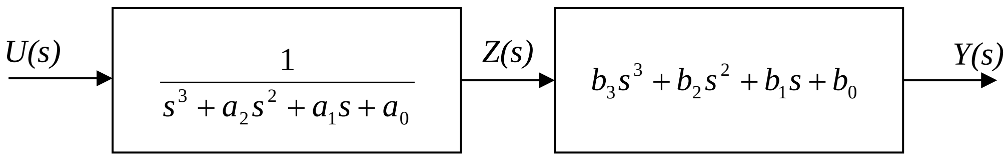
4.4.2 Transfer function → State space
Case 2: when the input does involve derivatives

- Consider the following third-order transfer function:

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

- Firstly, we split the transfer function into two parts by defining an intermediate variable *Z(s)* as follows:

$$\frac{Z(s)}{U(s)} = \frac{1}{s^3 + a_2 s^2 + a_1 s + a_0} \quad \text{and} \quad \frac{Y(s)}{Z(s)} = b_3 s^3 + b_2 s^2 + b_1 s + b_0$$



- Converting to the time domain gives:

$$\frac{d^3 z}{dt^3} + a_2 \frac{d^2 z}{dt^2} + a_1 \frac{dz}{dt} + a_0 z = u$$

and

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

- Setting the states as $x_1 = z, x_2 = \frac{dz}{dt}$ and $x_3 = \frac{d^2z}{dt^2}$, we get the state equation as follows:

$$\frac{dx_1}{dt} = x_2$$
$$\frac{dx_2}{dt} = x_3$$
$$\frac{dx_3}{dt} = -a_2x_3 - a_1x_2 - a_0x_1 + u$$

Once again, we have the **controllable canonical form**.

giving:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}u$$

- In terms of the output equation:

  - if $b_3 = 0$ then: $$y = b_2\frac{d^2z}{dt^2} + b_1\frac{dz}{dt} + b_0z = b_2x_3 + b_1x_2 + b_0x_1$$

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

    (*Note, the D matrix is 0 in this case*)

  - if $b_3 \neq 0$ then: $$y = b_2x_3 + b_1x_2 + b_0x_1 + b_3\frac{d^3z}{dt^3}$$

    but: $$\frac{d^3z}{dt^3} = \frac{dx_3}{dt} = -a_2x_3 - a_1x_2 - a_0x_1 + u$$

    giving:
    $$y = (b_2 - b_3a_2)x_3 + (b_1 - b_3a_1)x_2 + (b_0 - b_3a_0)x_1 + b_3u$$

therefore:

$$y = \left[(b_0 - b_3 a_0) \quad (b_1 - b_3 a_1) \quad (b_2 - b_3 a_2)\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + [b_3] u$$

- Note that the state matrix (i.e. matrix ***A***) is exactly the same for both types of transfer function models.

$$G(s) = \frac{Y(s)}{U(s)} = \frac{k}{s^3 + a_2 s^2 + a_1 s + a_0}$$

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

- This implies that the dynamics depend on the transfer function denominator only.

4.5 Obtaining transfer functions from state-space models

- In section 4.4, we saw how to go from a transfer function model to a state-space model (known as the **controllable canonical form**).

- It is also possible to go from a state-space model to a transfer function model as follows:

4.5.1 Continuous-time state-space model → transfer function

- Consider the following single-input-single-output continuous-time state-space model:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ y(t) &= \mathbf{C}\mathbf{x}(t) + \boldsymbol{D}u(t) \end{aligned}$$

- Taking the Laplace transform gives:

$$\begin{aligned} s\mathbf{X}(s) &= \mathbf{A}\mathbf{X}(s) + \mathbf{B}U(s) \\ Y(s) &= \mathbf{C}\mathbf{X}(s) + \mathbf{D}U(s) \end{aligned}$$

- Rearranging the state equation gives:

$$(sI - \boldsymbol{A})\boldsymbol{X}(s) = \boldsymbol{B}U(s)$$

$$\Rightarrow \boldsymbol{X}(s) = (sI - \boldsymbol{A})^{-1}\boldsymbol{B}U(s)$$

- Substituting this equation into the output equation gives:

$$Y(s) = \mathbf{C}(sI - \mathbf{A})^{-1}\mathbf{B}U(s) + \mathbf{D}U(s)$$

- Hence, the transfer function is defined in terms of the state-space equation matrices as:

$$G(s) = \frac{Y(s)}{U(s)} = \mathbf{C}(sI - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$$

- **Example 4.2:** Determine the transfer function for the following system:

$$\dot{\boldsymbol{x}}(t) = \begin{bmatrix} 0 & 2 \\ -3 & 5 \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t), \qquad y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}(t)$$

Solution

$$G(s) = \mathbf{C}(s\boldsymbol{I} - \mathbf{A})^{-1}\mathbf{B}$$

$$= \begin{bmatrix} 1 & 0 \end{bmatrix} \left(\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ -3 & 5 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s & -2 \\ 3 & s-5 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \frac{\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s-5 & 2 \\ -3 & s \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{s^2 - 5s + 6}$$

$$= \frac{\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s-5 \\ -3 \end{bmatrix}}{s^2 - 5s + 6} = \frac{s-5}{s^2 - 5s + 6}$$

Hence: $G(s) = \dfrac{Y(s)}{U(s)} = \dfrac{s-5}{s^2 - 5s + 6}$

l