Design of control systems involves the following three steps: modelling, analysis and design. Modelling involves finding a mathematical model of the plant (or system) being controlled. Analysis involves determining the behaviour of the system, e.g., stability, response of the system to reference inputs, etc. Design involves determining a controller which when used in the feedback loop with the system gives desired closed-loop system behaviour.

Modelling

We have studied two types of models in second and third years. They are:

- Transfer functions
- State-space Equations

Transfer Function

The transfer function of a system is defined as the ratio of Laplace transform of the output variable to the Laplace transform of the input variable.
System - Input \( u(t) \) → Output \( y(t) \)

\[
G(s) = \frac{\mathcal{L}\{y(t)\}}{\mathcal{L}\{u(t)\}} = \frac{Y(s)}{U(s)}
\]

State-Space Equations

\[
\begin{align*}
\dot{x}(t) &= A(t)x(t) + B(t)u(t) \\
y(t) &= C(t)x(t) + D(t)u(t)
\end{align*}
\]

If the linear system is also time invariant, then it is called a linear time-invariant (LTI) system

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t)
\end{align*}
\]

Remarks:

- Transfer function is a frequency domain model while the state space equations are time domain models.
- Transfer function model is unique for a system where as state space models are nonunique.
- The state-space equations gives information regarding the internal description in addition to the external description of the system where as the transfer function gives information regarding the external description.

**Derivation of Transfer Function from State Space Equations**

Consider a linear time-invariant system described by the following state-space equations:

\[
\begin{align*}
\dot{x}(t) & = Ax(t) + Bu(t) \\
y(t) & = Cx(t) + Du(t)
\end{align*}
\]

Assuming zero initial conditions and taking Laplace transform of the above state-space equations we have

\[
\begin{align*}
sX(s) & = AX(s) + BU(s) \\
Y(s) & = CX(s) + DU(s)
\end{align*}
\]

Simplifying the first equation give us

\[
(sI - A)X(s) = BU(s)
\]

or

\[
X(s) = (sI - A)^{-1}BU(s)
\]

Substituting the above in the output equation, we have

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

Therefore the transfer function is given by

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

The transfer function and state space equations are related by the following equation:

\[
G(s) = C(sI - A)^{-1}B + D
\]
Some definitions
(Ogata, 1997, Ref. [2])

**State** The state of a dynamic system is the smallest set of variables (called state variables) such that the knowledge of these variables at $t = t_0$ together with the knowledge of the input for $t \geq 0$ completely determines the behaviour of the system for any time $t \geq t_0$.

**State variables** The state variables of a dynamic system are the variables making up the smallest set of variables which determine the state of a dynamic system. If at least $n$ variables $x_1, x_2, \ldots, x_n$ are needed to completely describe the behaviour of a dynamic system, then such $n$ variables are called the state variables.

**State vector** If $n$ variables are needed to completely describe the behaviour of a given system then these $n$ state variables can be considered as the components of a vector $x$. Such a vector is called the state vector.

**State space** The $n$-dimensional state space whose co-ordinate axes consist of $x_1$ axis, $x_2$ axis, $\ldots$, $x_n$ axis is called the state space. Any state can be represented by a point in the state space.

**Example 1:** Consider a 3rd order system described by the differential equation:

$$
\begin{align*}
(3) \quad y' + 6 (2) \quad y'' + 11 (1) \quad y' + 6y &= 6u
\end{align*}
$$

where $u$ is the input and $y$ is the output. (i) Derive the state-space equations and (ii) compute the transfer function of the system.

**Example 2:** Derive the state-space equations for RLC network (Chen, 1984, Ref. [4], figure 3.4) shown in the figure:
Example 3: Derive the state-space equations using signal flow graphs for the following system:

\[ G(s) = \frac{b_1 s^2 + b_2 s + b_3}{s^3 + a_1 s^2 + a_2 s + a_3} \]

Multi-Input, Multi-Output (MIMO) System (see your text book)

A MIMO system has more than 1 input and output. A general MIMO system can be represented by the following state-space equations:

\[ \dot{x} = Ax + Bu \]
\[ y = Cx + Du \]

where \( A, B, C, \) and \( D \) are \( n \times n, n \times p, q \times n \) and \( q \times p \) matrices. The input \( u \) is a column vector of \( p \) elements and \( y \) is column vector of \( q \) elements. Expanding the state-space equations above, we can write:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\vdots \\
\dot{x}_n
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{12} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
+ \begin{bmatrix}
b_{11} & b_{12} & \cdots & b_{1p} \\
b_{12} & b_{22} & \cdots & b_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
b_{n1} & b_{n2} & \cdots & b_{np}
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2 \\
\vdots \\
u_p
\end{bmatrix}
\]

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_q
\end{bmatrix} =
\begin{bmatrix}
c_{11} & c_{12} & \cdots & c_{1n} \\
c_{12} & c_{22} & \cdots & c_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
c_{q1} & c_{q2} & \cdots & c_{qn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
+ \begin{bmatrix}
d_{11} & d_{12} & \cdots & d_{1p} \\
d_{12} & d_{22} & \cdots & d_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
d_{q1} & d_{q2} & \cdots & d_{qp}
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2 \\
\vdots \\
u_p
\end{bmatrix}
\]

The transfer function of a MIMO system is a \( q \times p \) matrix given by

\[ G(s) = C(sI - A)^{-1}B + D \]
\[
\begin{bmatrix}
T_{11}(s) & T_{12}(s) & \ldots & T_{1p} \\
T_{21}(s) & T_{22}(s) & \ldots & T_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
T_{q1}(s) & T_{q2}(s) & \ldots & T_{qp}
\end{bmatrix}
\]

where

\[T_{ij} = \frac{Y_i(s)}{U_j(s)}\]

is a scalar transfer function.

**Example 4:** Compute the transfer function matrices of the following systems:

1. 
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
-2 & 3 \\
-1 & -1
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
4 & 0 \\
-5 & 6
\end{bmatrix} \begin{bmatrix}
u_1 \\
u_2
\end{bmatrix} \]

\[y = \begin{bmatrix}
7 & 8
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}\]

2. 
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
-3 & 4 \\
-2 & 0
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
2 \\
u_1
\end{bmatrix} \]

\[
\begin{bmatrix}
y_1 \\
y_2
\end{bmatrix} = \begin{bmatrix}
-4 & 6 \\
5 & -1
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}\]

3. 
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
-2 & 3 \\
-1 & -1
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
4 & 0 \\
-5 & 6
\end{bmatrix} \begin{bmatrix}
u_1 \\
u_2
\end{bmatrix} \]

\[
\begin{bmatrix}
y_1 \\
y_2
\end{bmatrix} = \begin{bmatrix}
7 & 8 \\
5 & 6
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}\]
State-space realizations

Consider a 3rd order system described by the following transfer function:

\[ G(s) = \frac{b_1 s^2 + b_2 s + b_3}{s^3 + a_1 s^2 + a_2 s + a_3} \]

As state-space equations are non-unique, we get infinity of state-space models for the system, \( G(s) \). In this unit, we study some of the standard realizations (Kailath, 1980, Ref. [1]), we use often in the analysis and design of control systems.

- Controller canonical form
- Controllability canonical form
- Observer canonical form
- Observability canonical form
- Diagonal form or realization

**Controller canonical form**

This realization is described by the following state-space equations:

\[
\begin{align*}
\dot{x}_c &= A_c x_c + b_c u \\
y &= c_c x_c
\end{align*}
\]

where

\[
A_c = \begin{bmatrix}
-a_1 & -a_2 & -a_3 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{bmatrix},
\quad
b_c = \begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix},
\quad
\text{and } c_c = \begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix}
\]

**Observer canonical form**

This realization is described by the following state-space equations:

\[
\begin{align*}
\dot{x}_o &= A_o x_o + b_o u \\
y &= c_o x_o
\end{align*}
\]
where
\[
A_o = \begin{bmatrix}
-a_1 & 1 & 0 \\
-a_2 & 0 & 1 \\
-a_3 & 0 & 0
\end{bmatrix},
\quad b_o = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},
\quad c_o = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}
\]

Controllability canonical form

This realization is described by the following state-space equations:

\[
\begin{align*}
\dot{x}_{co} &= A_{co}x_{co} + b_{co}u \\
y &= c_{co}x_{co}
\end{align*}
\]

where
\[
A_{co} = \begin{bmatrix}
0 & 0 & -a_3 \\
1 & 0 & -a_2 \\
0 & 1 & -a_1
\end{bmatrix},
\quad b_{co} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}
\]

\[
c_{co} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \begin{bmatrix} 1 & a_1 & a_2 \\
0 & 1 & a_1 \\
0 & 0 & 1 \end{bmatrix}^{-1}
\]

Observability canonical form

This realization is described by the following state-space equations:

\[
\begin{align*}
\dot{x}_{ob} &= A_{ob}x_{ob} + b_{ob}u \\
y &= c_{ob}x_{ob}
\end{align*}
\]

where
\[
A_{ob} = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
-a_3 & -a_2 & -a_1
\end{bmatrix},
\quad c_{ob} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}
\]

\[
b_{ob} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\
1 & 0 & 0 \\
a_1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}
\]

b_{ob} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\
1 & 0 & 0 \\
a_1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}
Comments on Canonical forms

Suppose \{A, b, c\} is a realization of the system, \(G(s)\) then \(\{A^T, c^T, b^T\}\) is also a realization of \(G(s)\).

**Note:** \{\(A_c, b_c, c_c\)\} and \{\(A_o, b_o, c_o\)\} are dual realizations. Similarly \{\(A_{co}, b_{co}, c_{co}\)\} and \{\(A_{ob}, b_{ob}, c_{ob}\)\} are dual realizations.

**Theorem** In single-input, single output (SISO) case, dual realizations have the same transfer function.

**Proof:**

\[
G(s) = c(sI - A)^{-1}b
\]

For SISO case, \(G(s)\) is a scalar. Therefore, \(G(s) = G(s)^T\).

\[
G(s) = c(sI - A)^{-1}b = [c(sI - A)^{-1}b]^T = b^T(sI - A)^{-T}c^T = G(s)^T
\]

**Diagonal Realization**

In this section, we consider how to obtain a diagonal (or parallel) realization from a transfer function. We will assume that the system has only distinct poles.

Given a transfer function of the form:

\[
G(s) = \frac{\sum_{i=0}^{m} b_is^{m-i}}{\sum_{i=0}^{n} a_is^{n-i}} = \frac{b(s)}{a(s)}
\]

it can be expanded by partial fraction expansion as follows:

\[
G(s) = G_1(s) + G_2(s) + \ldots + G_n(s)
\]

where

\[
G_i(s) = \frac{g_i}{s - \lambda_i} \text{ and } g_i = b_ic_i
\]

This gives the following realization for \(G(s)\)
If the above block diagram is redrawn in time domain, the frequency domain signals \( U(s) \) and \( Y(s) \) can be represented by time domain signals \( u(t) \) and \( y(t) \) respectively, and the frequency domain block \( G_i(s) \) can be represented by the following time domain block diagram:

The following diagonal realization can be easily obtained from the time domain block diagram.

\[
\begin{bmatrix}
\dot{x}_1(t) \\
\dot{x}_2(t) \\
\vdots \\
\dot{x}_n(t)
\end{bmatrix} =
\begin{bmatrix}
\lambda_1 & 0 & 0 \\
0 & \lambda_1 & 0 \\
\vdots & \vdots & \vdots \\
0 & 0 & \lambda_n
\end{bmatrix}
\begin{bmatrix}
x_1(t) \\
x_2(t) \\
\vdots \\
x_n(t)
\end{bmatrix} +
\begin{bmatrix}
b_1 \\
b_2 \\
\vdots \\
b_n
\end{bmatrix} u
\]

\[
y(t) =
\begin{bmatrix}
c_1 & c_2 & \ldots & c_n
\end{bmatrix}
\begin{bmatrix}
x_1(t) \\
x_2(t) \\
\vdots \\
x_n(t)
\end{bmatrix}
\]
Example 5: Find a parallel/diagonal realization for the following transfer function:

\[ G(s) = \frac{2s^2 + 9s + 20}{s^3 + 6s^2 + 11s + 6} \]

Example 6: Find a realization for the following transfer function:

\[ G(s) = \frac{as + b}{cs^2 + ds + e} \]

where \( G(s) \) has complex conjugate poles.

Example 7: Find a parallel/diagonal realization for the following transfer function:

\[ G(s) = \frac{2s^2 + 6s + 5}{s^3 + 4s^2 + 5s + 2} \]

Example 8: Obtain a diagonal realization for the following transfer function matrices:

(a) \[ G(s) = \begin{bmatrix} \frac{s + 3}{s^2 + 3s + 2} & \frac{s + 4}{s^2 + 3s + 2} \\ \end{bmatrix} \]

(b) \[ G(s) = \begin{bmatrix} \frac{s + 5}{s^2 + 3s + 2} \\ \frac{s + 4}{s^2 + 3s + 2} \end{bmatrix} \]

**Markov Parameters (Kailath, 1980, Ref. [1])**

Markov parameters of a system \( G(s) = c(sI - A)^{-1}b \) are defined as follows:

\[ m_i = cA^{i-1}b, \quad i = 1, 2, \ldots \]

They are system invariants which means they are independent of similarity transformation and are dependent only on the transfer function of the system. Markov parameters are related to the impulse response of the system and they can shown to
be

\[ m_1 = g(t) \big|_{t=0} \]
\[ m_2 = \frac{dg(t)}{dt} \bigg|_{t=0} \]
\[ m_3 = \frac{d^2g(t)}{dt^2} \bigg|_{t=0} \]
\[ \vdots \]

where

\[ g(t) = \mathcal{L}^{-1} \{G(s)\} \]

Consider the expansion of the transfer function in the negative powers of \( s \) as shown below:

\[ G(s) = c(sI - A)b \]
\[ = c \left[ s(I - \frac{A}{s}) \right]^{-1} b \]
\[ = c \left[ \frac{1}{s}(I - \frac{A}{s})^{-1} \right] b \]
\[ = c \left[ \frac{I}{s} + \frac{A}{s^2} + \frac{A^2}{s^3} + \ldots \right] b \]
\[ = \frac{cb}{s} + \frac{cAb}{s^2} + \frac{cA^2b}{s^3} + \ldots \]

One can obtain Markov parameters of the system by expanding its transfer function in the negative powers of \( s \). Note that the transfer function expansion can also be written as follows:

\[ G(s) = \frac{b_1s^{n-1} + b_2s^{n-2} + \ldots + b_{n-1}s + b_n}{s^n + a_1s^{n-1} + a_2s^{n-2} + \ldots + a_{n-1}s + a_n} \]
\[ = \frac{b_1}{s} + \frac{(b_2 - a_1b_1)}{s^2} + \frac{(b_3 - a_1b_2 - a_2b_1 + a_1^2b_1)}{s^3} + \ldots \]

Note that the \( \beta \)'s used in controllability and observability canonical forms are Markov parameters.
Some Definitions from Linear Algebra
(Anton, 1984, Ref. [7])

Matrices and Vectors

A $m \times n$ matrix $A$ consists of a collection of $mn$ elements or quantities, $a_{ij}$ ($i = 1, 2, \ldots, m; j = 1, 2, \ldots, n$) written in form of an array of elements with $m$ rows and $n$ columns:

$$
A = \begin{bmatrix}
ap_{11} & a_{12} & \cdots & a_{1n} 
a_{21} & a_{22} & \cdots & a_{2n} 
\vdots & \vdots & \ddots & \vdots 
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
$$

If a matrix consists of only one row, then it is called a row vector; if it consists of only one column, then is called a column vector. Row and column vectors are also referred to by simply vectors.

Multiplication of Matrices

Consider the multiplication of two matrices with $A$ an $m \times p$ matrix and $B$ a $p \times n$ matrix. The product $C = AB$ is an $m \times n$ matrix with

$$
c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}
$$

Note that $C$ has the same number of rows of $A$, and the same number of columns as $B$.

Transposition

Suppose $A$ is an $m \times n$ matrix, then the transpose of $A$ (written as $A^T$ or $A'$) is an $n \times m$ matrix defined by

$$
B = A^T
$$

where $b_{ij} = a_{ji}$. Some useful formulas involving transpose of product of matrices:

$$
(AB)^T = B^TA^T
$$
$$
(ABC)^T = C^TB^TA^T
$$
Eigenvalues of a $n \times n$ matrix $A$ are the roots of the characteristic equation:

$$|\lambda I - A| = 0$$

The eigenvalues are sometimes called characteristic values.

**Eigenvectors of a $n \times n$ matrix:** Any non zero vector $x_i$ such that

$$Ax_i = \lambda x_i$$

is said to be an eigenvector associated with an eigenvalue $\lambda_i$ of $A$ where $A$ is a $n \times n$ matrix.

The eigenvalues are obtained by determining the roots of the characteristic equation:

$$|\lambda I - A| = 0$$

The eigenvectors are obtained by solving homogeneous algebraic equations of the form:

$$[A - \lambda_i I] x_i = 0$$

**Example 9:** Determine the eigenvalues and eigenvectors of the following $3 \times 3$ matrix:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$$

**Answer:** The eigenvalues are $\lambda_1 = -1$, $\lambda_2 = -2$, and $\lambda_3 = -3$.

The eigenvectors are:

$$x_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}, \quad \text{and} \quad x_3 = \begin{bmatrix} -3 \\ -3 \\ 9 \end{bmatrix}.$$
Determinants

The determinant of a \( n \times n \) matrix \( A \) can be computed by multiplying the entries in any row (or column) by their cofactors and adding the resulting products; that is for each \( 1 \leq i \leq n \) and \( 1 \leq j \leq n \)

\[
\det(A) = |A| = \sum_{i=1}^{n} a_{ij} C_{ij}
\]

(cofactor expansion along the \( j \)-th column)

\[
\det(A) = |A| = \sum_{j=1}^{n} a_{ij} C_{ij}
\]

(cofactor expansion along the \( i \)-th row)

**Exercise:** Determine the determinant of the following matrix

\[
A = \begin{bmatrix}
3 & 1 & 0 \\
-2 & -4 & 3 \\
5 & 4 & -2
\end{bmatrix}
\]

**Answer:** \( \det(A) = -1 \).

**Minors and Cofactors:** If \( A \) is a square matrix, then the **minor of entry** \( a_{ij} \) is denoted by \( M_{ij} \) and is defined to be determinant of the submatrix that remains after the \( i \)-th row and \( j \)-th column are deleted from \( A \). The number \( (-1)^{i+j} M_{ij} \) is denoted by \( C_{ij} \) and is called the **cofactor of entry** \( a_{ij} \).

**Exercise:** Let

\[
A = \begin{bmatrix}
3 & 1 & -4 \\
2 & 5 & 6 \\
1 & 4 & 8
\end{bmatrix}
\]

Determine \( M_{11}, C_{11}, M_{32} \) and \( C_{32} \).

**Answer:** \( M_{11} = 16, C_{11} = 16, M_{32} = 26, C_{32} = -26, \)
Matrix of Cofactors and Adjoint

If A is any $n \times n$ matrix and $C_{ij}$ is the cofactor of $a_{ij}$, then the matrix

$$
\begin{bmatrix}
C_{11} & C_{12} & \cdots & C_{1n} \\
C_{21} & C_{22} & \cdots & C_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
C_{n1} & C_{n2} & \cdots & C_{nn}
\end{bmatrix}
$$

is called the matrix of cofactors of A. The transpose of this matrix is called the adjoint of A and is denoted by $\text{adj}(A)$.

Inverse of a Matrix

If A is an $n \times n$ matrix, then its inverse is given by the following formula:

$$
A^{-1} = \frac{1}{\det(A)} \text{adj}(A)
$$

Example 10: Compute the inverse of a $2 \times 2$ matrix:

$$
A = \begin{bmatrix}
a & b \\
c & d
\end{bmatrix}
$$

Answer:

$$
A^{-1} = \frac{1}{ad - bc} \begin{bmatrix}
d & -b \\
-c & a
\end{bmatrix}
$$

Cayley-Hamilton Theorem states that every square matrix satisfies its own characteristic equation. If

$$
|\lambda I - A| = \lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n = 0
$$

is the characteristic equation then from Cayley Hamilton theorem we have

$$
A^n + a_1A^{n-1} + \ldots + a_{n-1}A + a_nI = 0
$$

Subspace: A subspace $W$ of a vector space $V$ is called a subspace of $V$ if $W$ itself is a vector space under the vector addition and scalar multiplication defined on $V$. 
Example 11: If W is a set of one or more vectors from a vector space V, then W is a subspace of V if and only if the following conditions hold.

1. If u and v are vectors in W, then \( u + w \) is in W.

2. If k is any scalar and u is any vector in W, then Ku is in W.

Note that conditions 1 and 2 means W is closed under vector addition and scalar multiplication.

Example 12: Show that the set W of all \( 2 \times 2 \) matrices having zeros on the main diagonal is a subspace of the vector space \( M_{22} \) of all \( 2 \times 2 \) matrices.

Solution: Let

\[
A = \begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & b_{12} \\ b_{21} & 0 \end{bmatrix}
\]

be any two matrices in W and k any scalar then

\[
kA = \begin{bmatrix} 0 & ka_{12} \\ ka_{21} & 0 \end{bmatrix}
\]

and

\[
A + B = \begin{bmatrix} 0 & a_{12} + b_{12} \\ a_{21} + b_{21} & 0 \end{bmatrix}
\]

Since kA and \( A + B \) have zeros on the main diagonal, they lie in W. Thus W is a subspace of \( M_{22} \).

Span: If \( v_1, v_2, \ldots, v_r \) are vectors in a vector space V and if every vector in V is expressible as a linear combination of \( v_1, v_2, \ldots, v_r \) then we say that these vectors span V.

Example 13: Vectors \( i = (1, 0, 0), j = (0, 1, 0) \), and \( k = (0, 0, 1) \) span \( R^3 \) because every vector \( (a, b, c) \) in \( R^3 \) can be written as

\[
(a, b, c) = ai + bj + ck
\]

which is linear combination of \( i, j, \) and \( k \).
State-space Approach: Advantages

[Masten, 1995, Ref. 6, Lesson 2]

The state-space approach is more general than the Laplace or Fourier transform techniques used for analysis of differential equations in classical control. As a result, state space techniques are applicable to all systems that can be analyzed by transform techniques. In addition, they are also suitable for many systems that cannot be analyzed by these classical techniques. The advantages of using state-space approach can be summarized as follows:

- Linear time varying systems can be analyzed in essentially the same manner as time-invariant linear systems.
- Multi-input and multi-output (MIMO) systems can be analyzed in the same manner as single-input and single-output systems.
- High-order systems are analyzed in the same way as low-order systems.
- Analysis, design, and implementation using state space techniques (which involve mainly matrix manipulations) can be easily done using a digital computer.
- The state-space approach can also be applied to analysis and design of nonlinear and/or stochastic systems.
Classical Control Vs Modern Control
[Masten, 1995, Ref. 6, Lesson 2]

The classical control (SS310) is a frequency domain approach and use the transfer function of the system as the mathematical model. Analysis tools are mainly developed for SISO systems. These tools include Bode plots (frequency response), Nyquist and Routh stability tests, root locus, and as well gain and phase margins. The modern control (CE421) is a time domain approach and use state-space equations of the system as the mathematical model. Analysis and design tools are developed for MIMO systems.

The classical control tools are mainly suitable for analysis whereas the modern control tools can be used for both analysis and design.

The design methods in classical control involve usually trial and error process whereas modern control methods are more systematic and controllers are designed using formulas.

Note that modern control techniques have not replaced classical control techniques, and many real-world problems are solved by classical procedures.

Computation of System Response
Ogata, 1997 (Ref. [2])

In this discussion we will consider only systems which are linear, continuous and time-invariant.
Solution of homogeneous state equations

We first consider scalar differential equation:

\[ \dot{x}(t) = ax(t) \]  

(1)

we assume a solution \( x(t) \) of the form

\[ x(t) = b_0 + b_1 t + b_2 t^2 + \ldots \]  

(2)

By substituting this assumed solution into (1) we get

\[ b_1 + 2b_2 t + 3b_2 t^2 + \ldots = a \left( b_0 + b_1 t + b_2 t^2 + \ldots \right) \]  

(3)

By equating coefficients of equal powers of \( t \), we get

\[ b_1 = ab_0 \]
\[ b_2 = \frac{1}{2} ab_1 = \frac{1}{2} a^2 b_0 \]
\[ b_3 = \frac{1}{3} ab_2 = \frac{1}{3 \times 2} a^3 b_0 \]
\[ \vdots \]
\[ b_k = \frac{1}{k!} a^k b_0 \]

The value of \( b_0 \) is determined by substituting \( t = 0 \) in eqn. (2), or

\[ x(0) = b_0 \]

Hence the solution \( x(t) \) can be written as

\[ x(t) = \left( 1 + at + \frac{1}{2!} a^2 t^2 + \ldots + \frac{1}{k!} a^k t^k + \ldots \right) x(0) \]
\[ = e^{at} x(0) \]

Solution of vector-matrix differential equation

We now consider vector-matrix differential equation:

\[ \dot{x}(t) = Ax(t) \]  

(4)
By analogy with the scalar case we assume a solution $x(t)$ is of the form of vector power series in $t$,

$$\quad
x(t) = B_0 + B_1 t + B_2 t^2 + \ldots \quad (5)$$

By substituting this assumed solution into (4) we get

$$B_1 + 2B_2 t + 3B_2 t^2 + \ldots = A (B_0 + B_1 t + B_2 t^2 + \ldots)$$

By equating coefficients of equal powers of $t$, we get

$$B_1 = AB_0$$

$$B_2 = \frac{1}{2} AB_1 = \frac{1}{2} A^2 B_0$$

$$B_3 = \frac{1}{3} AB_2 = \frac{1}{3 \times 2} A^3 B_0$$

$$\vdots$$

$$B_k = \frac{1}{k!} A^k B_0$$

The value of $B_0$ is determined by substituting $t = 0$ in eqn. (5), or

$$x(0) = B_0$$

Hence the solution $x(t)$ can be written as

$$x(t) = \left( I + At + \frac{1}{2!} A^2 t^2 + \ldots + \frac{1}{k!} A^k t^k + \ldots \right) x(0)$$

$$\quad = e^{At} x(0)$$

where

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}$$

The infinite series converges for all finite $t$. The matrix

$$\phi(t) = e^{At}$$

is known as the state transition matrix and has the following properties:
1. $\phi(0) = I$.

2. $\phi(t) = [\phi(-t)]^{-1}$ or $\phi^{-1}(t) = \phi(-t)$.

3. $\phi(t_1 + t_2) = \phi(t_1)\phi(t_2) = \phi(t_2)\phi(t_1)$.

4. $[\phi(t)]^n = \phi(nt)$.

5. $\phi(t_2 - t_1)\phi(t_1 - t_0) = \phi(t_2 - t_0)$.

6. $\frac{d}{dt} (e^{At}) = e^{At}A = Ae^{At}$.

**Exercise:** Prove the above properties of the state-transition matrix.

**Exercise:** Show that

$$e^{At} = \mathcal{L}^{-1} \left\{ (sI - A)^{-1} \right\}$$

### Solution of non-homogeneous state equation

We first consider the scalar case

$$\dot{x}(t) = ax(t) + bu(t)$$

rewriting the above equation

$$\dot{x}(t) - ax(t) = bu(t)$$

multiplying both sides by $e^{-at}$, we obtain

$$e^{-at}[\dot{x}(t) - ax(t)] = \frac{d}{dt}[e^{-at}x(t)] = e^{-at}bu(t)$$

integrating this equation between 0 and $t$, gives

$$e^{-at}x(t) = x(0) + \int_0^t e^{-a\tau}bu(\tau)d\tau$$

or

$$x(t) = e^{at}x(0) + e^{at} \int_0^t e^{-a\tau}bu(\tau)d\tau$$
we will now consider the vector-matrix non-homogeneous state equation described by

\[ \dot{x}(t) = Ax(t) + Bu(t) \]

where \( x(t) \) is a \( n \)-state vector, \( A \) is a \( n \times n \) matrix and \( B \) is a \( n \times 1 \) vector. By rewriting the above equation we get

\[ \dot{x}(t) - Ax(t) = Bu(t) \]

premultiplying both sides by \( e^{-At} \), we obtain

\[ e^{-At}[\dot{x}(t) - Ax(t)] = \frac{d}{dt}[e^{-At}x(t)] = e^{-At}Bu(t) \]

integrating this equation between 0 and \( t \), gives

\[ e^{-At}x(t) = x(0) + \int_0^t e^{-A\tau}Bu(\tau)d\tau \]

or

\[ x(t) = e^{At}x(0) + e^{At} \int_0^t e^{-A\tau}Bu(\tau)d\tau \]

The output of the system is given by

\[ y(t) = Cx(t) = Ce^{At}x(0) + Ce^{At} \int_0^t e^{-A\tau}Bu(\tau)d\tau \]

**Exercise:** Derive the above solutions using Laplace transform approach.

**Example 14:** For a second-order system given by the realization:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} =
\begin{bmatrix}
-3 & -2 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} +
\begin{bmatrix}
1 \\
0
\end{bmatrix} u
\]

\[ y = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \]

(i) Compute the state and output response for an impulse input assuming zero initial conditions (ii) Compute the state and output response for a step input and a non-zero initial condition given by \( x(0) = [1 \ 1]^T \).
Non-Uniqueness of Set of State Variables
Ogata, 1997 (Ref. [2])

State variables are not unique. For a given system, one can choose an infinite set of state variables to write state equations for the system. Suppose that \( x_1, x_2, \ldots, x_n \) are a set of state variables. Thus we may take as another set of state variables, any set of functions

\[
\begin{align*}
\dot{x}_1 &= X_1(x_1, x_2, \ldots, x_n) \\
\dot{x}_2 &= X_2(x_1, x_2, \ldots, x_n) \\
& \quad \vdots \\
\dot{x}_n &= X_n(x_1, x_2, \ldots, x_n)
\end{align*}
\]

provided that for every set of values \( \dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n \), there corresponds a unique set of values \( x_1, x_2, \ldots, x_n \) and vice versa. Thus, if \( x \) is a state vector, then \( \dot{x} \) given by

\[
\dot{x} = Px
\]

is also a state vector and vice versa.

**Similarity Transformation**

Consider a realization

\[
\begin{align*}
\dot{x} &= Ax + bu \quad (6) \\
y &= cx \quad (7)
\end{align*}
\]

of a system, \( G(s) \). Let

\[
\dot{x} = Px \quad \text{or} \quad x = P^{-1}\dot{x}
\]

where \( P \) is nonsingular. Substituting the above transformation in eqn. (6)-(7), we get

\[
\begin{align*}
\dot{x} &= P\dot{x} = PAx + Pbu \\
&= PAP^{-1}\dot{x} + Pbu
\end{align*}
\]
and

\[ y = cx = cP^{-1}\hat{x} \]

The new realization is given by

\[ \dot{\hat{x}} = \hat{A}\hat{x} + \hat{b}u \]
\[ y = \hat{c}\hat{x} \]

where

\[ \hat{A} = PAP^{-1} \]
\[ \hat{b} = Pb \]
\[ \hat{c} = cP^{-1} \]

If we choose the transformation as

\[ x = P\hat{x} \]

then the new realization will be

\[ \dot{\hat{x}} = \hat{A}\hat{x} + \hat{b}u \]
\[ y = \hat{c}\hat{x} \]

where

\[ \hat{A} = P^{-1}AP \]
\[ \hat{b} = P^{-1}b \]
\[ \hat{c} = cP \]

The realizations \( \{A, b, c\} \) and \( \{\hat{A}, \hat{b}, \hat{c}\} \) are similar realizations i.e., they are connected by similarity transformation.

**Properties of similar realizations**

1. Similar realizations have the same transfer function

Let

\[ G(s) = c(sI - A)^{-1}b \]
be the transfer function of the realization \( \{A, b, c\} \) and let
\[
\hat{G}(s) = \hat{c}(sI - \hat{A})^{-1}\hat{b}
\]
be the transfer function of the realization \( \{\hat{A}, \hat{b}, \hat{c}\} \). We will now show that
\[
G(s) = \hat{G}(s)
\]

**Proof:**
\[
\hat{G}(s) = \hat{c}(sI - \hat{A})^{-1}\hat{b}
\]
\[
= cP(sI - P^{-1}AP)^{-1}P^{-1}b
\]
\[
= c \left[ P(sI - P^{-1}AP)P^{-1} \right]^{-1}b
\]
\[
= c \left[ sIPP^{-1} - PP^{-1}APP^{-1} \right]^{-1}b
\]
\[
= c(sI - A)^{-1}b
\]

2. **Similar realizations have the same eigenvalues**

Recall that the eigenvalues of a matrix \( A \) are the roots of its characteristic equation:
\[
|\lambda I - A| = 0
\]

To show that similar realizations have the same eigenvalues, we have to show that they have the same characteristic equation, i.e.,
\[
|\lambda I - \hat{A}| = |\lambda I - A|
\]

**Proof:**
\[
|\lambda I - \hat{A}| = |\lambda I - P^{-1}AP|
\]
\[
= |P^{-1}(\lambda I - A)P|
\]
\[
= |P^{-1}|\lambda I - A||P|
\]
\[
= |P^{-1}||P||\lambda I - A|
\]
\[
= |\lambda I - A|
\]

**Question:** What is the relation between the poles of the system \( G(s) = c(sI - A)^{-1}b \) and the eigenvalues of the system matrix \( A \)?
Diagonalization

Consider a $n$-th order multi-input, and multi-output system described by

\[
\dot{x} = Ax + Bu \tag{8}
\]
\[
y = Cx + Du \tag{9}
\]

Assume that $A$ is non-diagonal. Let us define a new state vector such that

\[
x = Mz \tag{10}
\]

where $M$ is nonsingular. Substituting (3) in (1) and (2) yields

\[
\dot{z} = \hat{A}z + \hat{B}u
\]
\[
y = \hat{C}z + Du
\]

where

\[
\hat{A} = M^{-1}AM \\
\hat{B} = M^{-1}B \\
\hat{C} = CM
\]

If $M$ can be selected such that $M^{-1}AM$ is diagonal, then $M$ is called the diagonalizing matrix or the modal matrix. If matrix $A$ has $n$ linearly independent eigenvectors then the eigenvector matrix of $A$ can be chosen as the modal matrix.

**Remark:** If matrix $A$ has $n$ distinct eigenvalues then the matrix $A$ has $n$ linearly independent eigenvectors.

If a $n \times n$ matrix $A$ with distinct eigenvalues is given by

\[
A = \begin{bmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \vdots & & 1 \\
-a_n & -a_{n-1} & -a_{n-2} & -a_1
\end{bmatrix}
\]
then the modal matrix or the eigenvector matrix takes a special form given by
\[
P = \begin{bmatrix}
1 & 1 & \cdots & 1 \\
\lambda_1 & \lambda_2 & \cdots & \lambda_n \\
\lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1}
\end{bmatrix}
\]

The diagonal matrix is given by
\[
\hat{A} = P^{-1}AP = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}
\]

If matrix \(A\) has multiple eigenvalues then diagonalization is impossible because \(P^{-1}\) does not exist. In such cases the transformation matrix of the form:
\[
P = \begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
\lambda_1 & 1 & 0 & \cdots & 0 \\
\lambda_1^2 & 2\lambda_1 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\lambda_1^{n-1} & d\lambda_1^{n-1} & \frac{d^2}{d\lambda_1^2}\lambda_1^{n-1} & \cdots & 1
\end{bmatrix}
\]
yields a diagonalized matrix of the following form:
\[
\hat{A} = P^{-1}AP = \begin{bmatrix}
\lambda_1 & 1 & 0 & 0 \\
0 & \lambda_1 & 1 & 0 \\
\vdots & \vdots & \ddots & 1 \\
0 & 0 & 0 & \lambda_1
\end{bmatrix}
\]

This form is known as the Jordan form.

**Question:** What is the relation between the realization obtained by partial fraction expansion and the realization obtained by diagonalization?

**Example 15:** Diagonalise the following realization:
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3
\end{bmatrix} = \begin{bmatrix}
-1 & -2 & 0 \\
1 & 2 & 0 \\
-2 & -1 & -3
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix} + \begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix} u
\]
\[
\begin{bmatrix}
y_1 \\
y_2
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 1 \\
1 & -1 & 0
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
\]