Modeling General Concepts

• Basic building blocks of lumped-parameter modeling of real systems
  – mechanical
  – electrical
  – fluid
  – thermal
  – mixed with energy-conversion devices

• Real devices are modeled as combinations of pure and ideal elements
• There are several general concepts and methods of system dynamics that apply to all kinds of elements and systems.

• These concepts are of general applicability and we will address some of these now:
  – Classification of System Inputs
  – Pure and Ideal Elements vs. Real Devices
  – Ideal vs. Real Sources
  – Transfer Functions
  – Step Response
– Frequency Response
– Linearization of Nonlinear Physical Effects
– Block Diagram
– Loading Effects
– State-Space Representation
– Poles and Zeros of Transfer Functions
– Sensitivity Analysis
Input / System / Output Concept: Classification of System Inputs
Classification of System Inputs

- **Input** – some agency which can cause a system to respond.
- **Initial energy storage** refers to a situation in which a system, at time $= 0$, is put into a state different from some reference equilibrium state and then released, free of external driving agencies, to respond in its characteristic way. Initial energy storage can take the form of either kinetic energy or potential energy.
- **External driving agencies** are physical quantities which vary with time and pass from the external environment, through the system interface or boundary, into the system, and cause it to respond.
- We often choose to study the system response to an assumed ideal source, which is unaffected by the system to which it is coupled, with the view that practical situations will closely correspond to this idealized model.

- External inputs can be broadly classified as deterministic or random, recognizing that there is always some element of randomness and unpredictability in all real-world inputs.

- Deterministic input models are those whose complete time history is explicitly given, as by mathematical formula or a table of numerical values. This can be further divided into:
  - transient input model: one having any desired shape, but existing only for a certain time interval, being constant before the beginning of the interval and after its end.
– periodic input model: one that repeats a certain wave form over and over, ideally forever, and is further classified as either sinusoidal or non-sinusoidal.

– almost periodic input model: continuing functions which are completely predictable but do not exhibit a strict periodicity, e.g., amplitude-modulated input.

• Random input models are the most realistic input models and have time histories which cannot be predicted before the input actually occurs, although statistical properties of the input can be specified.

• When working with random inputs, there is never any hope of predicting a specific time history before it occurs, but statistical predictions can be made that have practical usefulness.
• If the statistical properties are time-invariant, then the input is called a **stationary random input**. **Unstationary random inputs** have time-varying statistical properties. These are often modeled as stationary over restricted periods of time.
Pure and Ideal Elements vs. Real Devices

- A pure element refers to an element (spring, damper, inertia, resistor, capacitor, inductor, etc.) which has only the named attribute.
- For example, a pure spring element has no inertia or friction and is thus a mathematical model (approximation), not a real device.
- The term ideal, as applied to elements, means linear, that is, the input/output relationship of the element is linear, or straight-line. The output is perfectly proportional to the input.
- A device can be pure without being ideal and ideal without being pure.
• From a functional engineering viewpoint, nonlinear behavior may often be preferable, even though it leads to difficult equations.

• Why do we choose to define and use pure and ideal elements when we know that they do not behave like the real devices used in designing systems? Once we have defined these pure and ideal elements, we can use these as building blocks to model real devices more accurately.

• For example, if a real spring has significant friction, we model it as a combination of pure/ideal spring and damper elements, which may come quite close in behavior to the real spring.
Ideal vs. Real Sources

- External driving agencies are physical quantities which pass from the environment, through the interface into the system, and cause the system to respond.
- In practical situations, there may be interactions between the environment and the system; however, we often use the concept of ideal source.
- An ideal source (force, motion, voltage, current, etc.) is totally unaffected by being coupled to the system it is driving.
- For example, a “real” 6-volt battery will not supply 6 volts to a circuit! The circuit will draw some current from the battery and the battery’s voltage will drop.
**Transfer Functions**

- **Definition and Comments**
  - The transfer function of a linear, time-invariant, differential equation system is defined as the ratio of the Laplace transform of the output (response function) to the Laplace transform of the input (driving function) under the assumption that all initial conditions are zero.
  - By using the concept of transfer function, it is possible to represent system dynamics by algebraic equations in $s$. The highest power of $s$ in the denominator determines the order of the system.
– The transfer function is a property of a system itself, independent of the magnitude and nature of the input or driving function.

– The transfer function gives a full description of the dynamic characteristics of the system.

– The transfer function does not provide any information concerning the physical structure of the system; the transfer functions of many physically different systems can be identical.

– If the transfer function of a system is known, the output or response can be studied for various forms of inputs with a view toward understanding the nature of the system.
If the transfer function of a system is unknown, it may be established experimentally by introducing known inputs and studying the output of the system.

**Convolution Integral**

- For a linear time-invariant system the transfer function $G(s)$ is
  \[ G(s) = \frac{Y(s)}{X(s)} \]

  where $X(s)$ is the Laplace transform of the input and $Y(s)$ is the Laplace transform of the output, assuming all initial conditions are zero.

- The inverse Laplace transform is given by the convolution integral:
  \[ y(t) = \int_0^t x(\tau)g(t - \tau)d\tau = \int_0^t g(\tau)x(t - \tau)d\tau \quad \text{for } t < 0 \]
  \[
  \begin{cases} 
  g(t) = 0 \\
  x(t) = 0
  \end{cases}
  \]
• **Impulse-Response Function**

  – The Laplace transform of the response of a system to a unit-impulse input, when the initial conditions are zero, is the transfer function of the system, i.e., \( Y(s) = G(s) \).

  – The inverse Laplace transform of the system transfer function, \( g(t) \), is called the impulse-response function. It is the response of a linear system to a unit-impulse response when the initial conditions are zero.

  – The transfer function and the impulse-response function of a linear, time-invariant system contain the same information about the system dynamics.
Experimentally, one can excite a system at rest with an impulse input (a pulse input of very short duration compared with the significant time constants of the system) and measure the response. This response is the impulse-response function, the Laplace transform of which is the transfer function of the system.
Step Response

- The step response of a system is the response of the system to a step input of some variable.
- By a step input of any variable, we will always mean a situation where the system is at rest at time $t = 0$ and we instantly change the input quantity, from wherever it was just before $t = 0$, by a given amount, either positive or negative, and then keep the input constant at this new value forever.
Frequency Response

- **Linear ODE with Constant Coefficients**

\[
a_n \frac{d^n q_o}{dt^n} + a_{n-1} \frac{d^{n-1} q_o}{dt^{n-1}} + \ldots + a_1 \frac{dq_o}{dt} + a_0 q_o = \\
b_m \frac{d^m q_i}{dt^m} + b_{m-1} \frac{d^{m-1} q_i}{dt^{m-1}} + \ldots + b_1 \frac{dq_i}{dt} + b_0 q_i
\]

- \(q_o\) is the output (response) variable of the system
- \(q_i\) is the input (excitation) variable of the system
- \(a_n\) and \(b_m\) are the physical parameters of the system
• If the input to a linear system is a sine wave, the steady-state output \((after\ the\ transients\ have\ died\ out)\) is also a sine wave with the same frequency, but with a different amplitude and phase angle.

• System Input: \(q_i = Q_i \sin(\omega t)\)

• System Steady-State Output: \(q_o = Q_o \sin(\omega t + \phi)\)

• Both amplitude ratio, \(Q_o/Q_i\), and phase angle, \(\phi\), change with frequency, \(\omega\).

• The frequency response can be determined analytically from the Laplace transfer function:

\[
G(s) \quad s = i\omega \quad \text{Sinusoidal Transfer Function} \quad M(\omega) \angle \phi(\omega)
\]
• A negative phase angle is called \textit{phase lag}, and a positive phase angle is called \textit{phase lead}.

• If the system being excited were a nonlinear or time-varying system, the output might contain frequencies other than the input frequency and the output-input ratio might be dependent on the input magnitude.

• Any real-world device or process will only need to function properly for a certain range of frequencies; outside this range we don’t care what happens.
Mechatronics
Modeling - General Concepts

System Frequency Response
(linear scales used)
Analog Electronics:  
RC Low-Pass Filter 
Time Response & Frequency Response 

\[
\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{1}{RCs + 1}
\]

Time Constant \( \tau = RC \)

\begin{align*}
\text{Resistor} & \quad 15 \text{ K}\Omega \\
\text{Capacitor} & \quad 0.01 \mu\text{F}
\end{align*}

Time Response
Frequency Response

\[
\frac{V_{\text{out}}(i\omega)}{V_{\text{in}}} = \frac{K}{i\omega \tau + 1} = \frac{K \angle 0^\circ}{\sqrt{(\omega \tau)^2 + 1^2} \angle \tan^{-1} \omega \tau} = \frac{K}{\sqrt{(\omega \tau)^2 + 1^2}} \angle -\tan^{-1} \omega \tau
\]

Bandwidth = \(1/\tau\)
MatLab / Simulink Diagram
Frequency Response for 1061 Hz Sine Input

\[ \tau = 1.5 \times 10^{-4} \text{ sec} \]

Sine Wave

First-Order Plant

\[ \frac{1}{\tau u + 1} \]

Clock

Time

Output
Amplitude Ratio = 0.707 = -3 dB
Phase Angle = -45°
• When one has the frequency-response curves for any system and is given a specific sinusoidal input, it is an easy calculation to get the sinusoidal output.

• What is not obvious, but extremely important, is that the frequency-response curves are really a complete description of the system’s dynamic behavior and allow one to compute the response for any input, not just sine waves.

• Every dynamic signal has a frequency spectrum and if we can compute this spectrum and properly combine it with the system’s frequency response, we can calculate the system time response.
• The details of this procedure depend on the nature of the input signal; is it periodic, transient, or random?

• For periodic signals (those that repeat themselves over and over in a definite cycle), Fourier Series is the mathematical tool needed to solve the response problem.

• Although a single sine wave is an adequate model of some real-world input signals, the generic periodic signal fits many more practical situations.

• A periodic function $q_i(t)$ can be represented by an infinite series of terms called a Fourier Series.
\[ q_i(t) = \frac{a_0}{T} + \frac{2}{T} \sum_{n=1}^{\infty} \left[ a_n \cos\left(\frac{2\pi n}{T} t\right) + b_n \sin\left(\frac{2\pi n}{T} t\right) \right] \]

\[
\begin{align*}
a_n &= \int_{-T/2}^{T/2} q_i(t) \cos\left(\frac{2\pi n}{T} t\right) dt \\
b_n &= \int_{-T/2}^{T/2} q_i(t) \sin\left(\frac{2\pi n}{T} t\right) dt
\end{align*}
\]

\textbf{Fourier Series}
Consider the Square Wave:

\[
\begin{align*}
\int_{-0.01}^{0} -0.5\,dt + \int_{0}^{0.01} 1.5\,dt &= 0.5 = \text{average value} \\
\int_{-0.01}^{0} -0.5\cos\left(\frac{2\pi n}{0.02} t\right)\,dt + \int_{0}^{0.01} 1.5\cos\left(\frac{2\pi n}{0.02} t\right)\,dt &= 0 \\
\int_{-0.01}^{0} -0.5\sin\left(\frac{2\pi n}{0.02} t\right)\,dt + \int_{0}^{0.01} 1.5\sin\left(\frac{2\pi n}{0.02} t\right)\,dt &= \frac{1 - \cos(n\pi)}{50n\pi} \\
q_i(t) &= 0.5 + \frac{4}{\pi}\sin(100\pi t) + \frac{4}{3\pi}\sin(300\pi t) + \ldots
\end{align*}
\]
• The term for \( n = 1 \) is called the fundamental or first harmonic and always has the same frequency as the repetition rate of the original periodic wave form (50 Hz in this example); whereas \( n = 2, 3, \ldots \) gives the second, third, and so forth harmonic frequencies as integer multiples of the first.

• The square wave has only the first, third, fifth, and so forth harmonics. The more terms used in the series, the better the fit. An infinite number gives a “perfect” fit.
Plot of the Fourier Series for the square wave through the third harmonic

\[ q_i(t) = 0.5 + \frac{4}{\pi} \sin(100\pi t) + \frac{4}{3\pi} \sin(300\pi t) \]
• For a signal of arbitrary periodic shape (rather than the simple and symmetrical square wave), the Fourier Series will generally include all the harmonics and both sine and cosine terms.

• We can combine the sine and cosine terms using:

\[ A \cos(\omega t) + B \sin(\omega t) = C \sin(\omega t + \alpha) \]

\[ C = \sqrt{A^2 + B^2} \]

\[ \alpha = \tan^{-1} \frac{A}{B} \]

• Thus

\[ q_i(t) = A_{i0} + A_{i1} \sin(\omega_1 t + \alpha_1) + A_{i2} \sin(2\omega_1 t + \alpha_2) + \ldots \]
• A graphical display of the amplitudes ($A_{ik}$) and the phase angles ($\alpha_k$) of the sine waves in the equation for $q_i(t)$ is called the frequency spectrum of $q_i(t)$.

• If a periodic $q_i(t)$ is applied as input to a system with sinusoidal transfer function $G(i\omega)$, after the transients have died out, the output $q_o(t)$ will be in a periodic steady-state given by:

$$q_o(t) = A_{o0} + A_{o1} \sin (\omega_1 t + \theta_1) + A_{o2} \sin (2\omega_1 t + \theta_2) + ...$$

$$A_{ok} = A_{ik} \left| G(i\omega_k) \right|$$

$$\theta_k = \alpha_k + \angle G(i\omega_k)$$

• This follows from superposition and the definition of the sinusoidal transfer function.
Linearization of Nonlinear Physical Effects

- Many real-world nonlinearities involve a "smooth" curvilinear relation between an independent variable $x$ and a dependent variable $y$:
  \[ y = f(x) \]

- A linear approximation to the curve, accurate in the neighborhood of a selected operating point, is the tangent line to the curve at this point.

- This approximation is given conveniently by the first two terms of the Taylor series expansion of $f(x)$:

\[
\begin{align*}
\hat{y} &= \bar{y} + \left. \frac{df}{dx} \right|_{x=\bar{x}} (x - \bar{x}) \\
\hat{y} &= \bar{y} + \frac{df}{dx} \left|_{x=\bar{x}} \right. (x - \bar{x}) \\
\hat{y} &= \bar{y} + \frac{df}{dx} \left|_{x=\bar{x}} \right. (x - \bar{x}) \\
\hat{y} &= K \hat{x}
\end{align*}
\]
• For example, in liquid-level control systems, when the tank is not prismatic, a nonlinear volume/height relationship exists and causes a nonlinear system differential equation. For a conical tank of height $H$ and top radius $R$ we would have:

$$V = \frac{\pi R^2}{3H^2} h^3$$

$$V \approx \frac{\pi R^2 h^3}{3H^2} + \frac{\pi R^2 h^2}{H^2} h$$

• Often a dependent variable $y$ is related nonlinearly to several independent variables $x_1$, $x_2$, $x_3$, etc. according to the relation: $y = f(x_1, x_2, x_3, \ldots)$. 
We may linearize this relation using the multivariable form of the Taylor series:

\[ y \approx f(\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots) + \left. \frac{\partial f}{\partial x_1} \right|_{\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots} (x_1 - \overline{x}_1) + \left. \frac{\partial f}{\partial x_2} \right|_{\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots} (x_2 - \overline{x}_2) \]

\[ + \left. \frac{\partial f}{\partial x_3} \right|_{\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots} (x_3 - \overline{x}_3) + \ldots \]

\[ y \approx \bar{y} + \left. \frac{\partial f}{\partial x_1} \right|_{\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots} \hat{x}_1 + \left. \frac{\partial f}{\partial x_2} \right|_{\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots} \hat{x}_2 + \left. \frac{\partial f}{\partial x_3} \right|_{\overline{x}_1, \overline{x}_2, \overline{x}_3, \ldots} \hat{x}_3 + \ldots \]

\[ \hat{y} = K_1 \hat{x}_1 + K_2 \hat{x}_2 + K_3 \hat{x}_3 + \ldots \]

The partial derivatives can be thought of as the sensitivity of the dependent variable to small changes in that independent variable.
For example, in a ported gas-filled piston/cylinder where gas mass, temperature, and volume are all changing, the perfect gas law gives us for pressure $p$:

$$p = \frac{RTM}{V}$$

$$p \approx \frac{RTM}{V} + \frac{RM}{V} (T - \bar{T}) + \frac{RT}{V} (M - \bar{M}) - \frac{RMT}{V} (V - \bar{V})$$
Example: Magnetic Levitation System
Electromagnet

\( f(x,i) = C \left( \frac{i^2}{x^2} \right) \)

Ball (mass m)

\[ mg \]

\[ m \ddot{x} = mg - C \left( \frac{i^2}{x^2} \right) \]

\[ m \dddot{x} = mg - C \left( \frac{i^2}{x^2} \right) + C \left( \frac{2i}{x^2} \right) \dot{x} - C \left( \frac{2i}{x^2} \right) i \]

**Magnetic Levitation System**

**Equation of Motion:**

\[ m \ddot{x} = mg - C \left( \frac{i^2}{x^2} \right) \]

**At Equilibrium:**

\[ mg = C \left( \frac{i^2}{x^2} \right) \]

**Linearization:**

\[ C \left( \frac{i^2}{x^2} \right) \approx C \left( \frac{i^2}{x^2} \right) - C \left( \frac{2i}{x^2} \right) \dot{x} + C \left( \frac{2i}{x^2} \right) i \]
Use of Experimental Testing in Multivariable Linearization

\[ f_m = f(i, y) \]

\[ f_m \approx f(i_0, y_0) + \left. \frac{\partial f}{\partial y} \right|_{i_0, y_0} (y - y_0) + \left. \frac{\partial f}{\partial i} \right|_{i_0, y_0} (i - i_0) \]
Block Diagrams

• A block diagram of a system is a pictorial representation of the functions performed by each component and of the flow of signals. It depicts the interrelationships that exist among the various components.

• It is easy to form the overall block diagram for the entire system by merely connecting the blocks of the components according to the signal flow. It is then possible to evaluate the contribution of each component to the overall system performance.

• A block diagram contains information concerning dynamic behavior, but it does not include any information on the physical construction of the system.
• Many dissimilar and unrelated systems can be represented by the same block diagram.
• A block diagram of a given system is not unique. A number of different block diagrams can be drawn for a system, depending on the point of view of the analysis.
• Closed-Loop System Block Diagram:
\[
\frac{B(s)}{E(s)} = G_{c}(s)G(s)H(s)
\]

Open-Loop Transfer Function

\[
\frac{C(s)}{E(s)} = G_{c}(s)G(s)
\]

Feedforward Transfer Function

Closed-Loop Transfer Functions

\[
\frac{C(s)}{R(s)} = \frac{G_{c}(s)G(s)}{1 + G_{c}(s)G(s)H(s)}
\]

\[
|G_{c}(s)G(s)H(s)| >> 1 \quad \Rightarrow \quad \frac{C(s)}{R(s)} \Rightarrow \frac{1}{H(s)}
\]

\[
\frac{C(s)}{D(s)} = \frac{G(s)}{1 + G_{c}(s)G(s)H(s)}
\]

\[
|G_{c}(s)G(s)H(s)| >> 1 \quad \Rightarrow \quad \frac{C(s)}{D(s)} \Rightarrow 0
\]

\[
C(s) = \frac{G(s)}{1 + G_{c}(s)G(s)H(s)} \left[ G_{c}(s)R(s) + D(s) \right]
\]
• Blocks can be connected in series only if the output of one block is not affected by the next following block. If there are any loading effects between components, it is necessary to combine these components into a single block.

• In simplifying a block diagram, remember:
  – The product of the transfer functions in the feedforward direction must remain the same.
  – The product of the transfer functions around the loop must remain the same.
### Rules of Block Diagram Algebra

<table>
<thead>
<tr>
<th>Original Block Diagrams</th>
<th>Equivalent Block Diagrams</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram 1" /></td>
<td><img src="image2" alt="Diagram 2" /></td>
</tr>
<tr>
<td><img src="image3" alt="Diagram 3" /></td>
<td><img src="image4" alt="Diagram 4" /></td>
</tr>
<tr>
<td><img src="image5" alt="Diagram 5" /></td>
<td><img src="image6" alt="Diagram 6" /></td>
</tr>
</tbody>
</table>
Loading Effects

• The unloaded transfer function is an incomplete component description.

• To properly account for interconnection effects one must know three component characteristics:
  – the unloaded transfer function of the upstream component
  – the output impedance of the upstream component
  – the input impedance of the downstream component

• Only when the ratio of output impedance over input impedance is small compared to 1.0, over the frequency range of interest, does the unloaded transfer function give an accurate description of interconnected system behavior.
\[
\begin{align*}
\frac{Y(s)}{U(s)} &= \begin{bmatrix}
G_1(s) - \frac{1}{Z_{o1}} \\
1 + \frac{Z_{o1}}{Z_{i2}}
\end{bmatrix} G_2(s)
\end{align*}
\]

Only if \( \frac{Z_{o1}}{Z_{i2}} \ll 1 \) for the frequency range of interest will loading effects be negligible
• In general, loading effects occur because when analyzing an isolated component (one with no other component connected at its output), we assume no power is being drawn at this output location.

• When we later decide to attach another component to the output of the first, this second component does withdraw some power, violating our earlier assumption and thereby invalidating the analysis (transfer function) based on this assumption.

• When we model chains of components by simple multiplication of their individual transfer functions, we assume that loading effects are either not present, have been proven negligible, or have been made negligible by the use of buffer amplifiers.
Analog Electronics
Example: Loading Effects

\[
\begin{bmatrix}
V_{in} \\
\mathbf{i}_{in}
\end{bmatrix} =
\begin{bmatrix}
RCs + 1 & -R \\
Cs & -1
\end{bmatrix}
\begin{bmatrix}
V_{out} \\
\mathbf{i}_{out}
\end{bmatrix}
\]

\[
\frac{V_{out}}{V_{in}} = \frac{1}{RCs + 1} = \frac{1}{\tau s + 1}
\]

when \( \mathbf{i}_{out} = 0 \) 

RC Low-Pass Filter

\[
Z_{out} = \left. \frac{V_{out}}{\mathbf{i}_{out}} \right|_{V_{in} = 0} = \frac{R}{RCs + 1}
\]

Output Impedance

\[
Z_{in} = \left. \frac{V_{in}}{\mathbf{i}_{in}} \right|_{\mathbf{i}_{out} = 0} = \frac{RCs + 1}{Cs}
\]

Input Impedance
2 RC Low-Pass Filters in Series

Only if $Z_{out-1} << Z_{in-2}$ for the frequency range of interest will loading effects be negligible.

\[
\frac{V_{out}}{V_{in}} \neq G(s)_{1-unloaded} G(s)_{2-unloaded} = \left( \frac{1}{RC_s + 1} \right) \left( \frac{1}{RC_s + 1} \right)
\]

\[
\frac{V_{out}}{V_{in}} = G(s)_{1-loaded} G(s)_{2-unloaded}
\]

\[
= \left( \frac{1}{RC_s + 1} \right) \left( 1 + \frac{1}{Z_{out-1}} \right) \left( \frac{1}{RC_s + 1} \right) = \frac{1}{(RC_s + 1)^2 + RC_s}
\]

Resistor 15 KΩ
Capacitor 0.01 µF

Resistor 15 KΩ
Capacitor 0.01 µF
State-Space Representation

- Conventional Control Theory (root-locus and frequency response analysis and design) is applicable to linear, time-invariant, single-input, single-output systems. This is a complex frequency-domain approach. The transfer function relates the input to output and does not show internal system behavior.

- Modern Control Theory (state-space analysis and design) is applicable to linear or nonlinear, time-varying or time-invariant, multiple-input, multiple-output systems. This is a time-domain approach. This state-space system description provides a complete internal description of the system, including the flow of internal energy.
A state-determined system is a special class of lumped-parameter dynamic system such that: (i) specification of a finite set of \( n \) independent parameters, state variables, at time \( t = t_0 \) and (ii) specification of the system inputs for all time \( t \geq t_0 \) are necessary and sufficient to uniquely determine the response of the system for all time \( t \geq t_0 \).

The state is the minimum amount of information needed about the system at time \( t_0 \) such that its future behavior can be determined without reference to any input before \( t_0 \).
- The *state variables* are independent variables capable of defining the state from which one can completely describe the system behavior. These variables completely describe the effect of the past history of the system on its response in the future.

- Choice of state variables is not unique and they are often, but not necessarily, physical variables of the system. They are usually related to the energy stored in each of the system's energy-storing elements, since any energy initially stored in these elements can affect the response of the system at a later time.
– State variables do not have to be physical or measurable quantities, but practically they should be chosen as such since optimal control laws will require the feedback of all state variables.

– The state-space is a conceptual n-dimensional space formed by the n components of the state vector. At any time \( t \) the state of the system may be described as a point in the state space and the time response as a trajectory in the state space.

– The number of elements in the state vector is unique, and is known as the order of the system.
– Since integrators in a continuous-time dynamic system serve as memory devices, the outputs of integrators can be considered as state variables that define the internal state of the dynamic system. Thus the outputs of integrators can serve as state variables.

– The state-variable equations are a coupled set of first-order ordinary differential equations where the derivative of each state variable is expressed as an algebraic function of state variables, inputs, and possibly time.
\[
\dot{x}(t) = \dot{f}(\dot{x}, \ddot{u}, t) \\
\ddot{y}(t) = \ddot{g}(\dot{x}, \ddot{u}, t) \\
\text{Non-Linear, Time-Varying}
\]

\[
\begin{align*}
\dot{x}(t) &= A(t)\dot{x}(t) + B(t)\ddot{u}(t) \\
\ddot{y}(t) &= C(t)\dot{x}(t) + D(t)\ddot{u}(t)
\end{align*}
\text{Linear, Time-Varying}
\]
State-Space to Transfer Function

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

\[
\frac{Y(s)}{U(s)} = G(s)
\]

Laplace Transform

\[
sX(s) - x(0) = AX(s) + BU(s)
\]

\[
Y(s) = CX(s) + DU(s)
\]

Zero Initial Conditions

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

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

Define:

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

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

\[
Y(s) = [C\Phi(s)B + D]U(s)
\]

\[
\frac{Y(s)}{U(s)} = [C\Phi(s)B + D] = G(s)
\]
• The poles of the transfer function are the eigenvalues of the system matrix A.
  \[ |sI - A| = 0 \]  
  Characteristic Equation

• A zero of the transfer function is a value of \( s \) that satisfies:
  \[ |sI - \begin{bmatrix} A & -B \\ C & D \end{bmatrix}| = 0 \]

• The transfer function can be written as:
  \[
  G(s) = \frac{|sI - \begin{bmatrix} A & -B \\ C & D \end{bmatrix}|}{|sI - A|}
  \]
Poles and Zeros of Transfer Functions

- Definition of Poles and Zeros
  - A pole of a transfer function $G(s)$ is a value of $s$ (real, imaginary, or complex) that makes the denominator of $G(s)$ equal to zero.
  - A zero of a transfer function $G(s)$ is a value of $s$ (real, imaginary, or complex) that makes the numerator of $G(s)$ equal to zero.
  - For Example: $G(s) = \frac{K(s + 2)(s + 10)}{s(s + 1)(s + 5)(s + 15)^2}$
    - Poles: 0, -1, -5, -15 (order 2)
    - Zeros: -2, -10, $\infty$ (order 3)
• **Collocated Control System**
  – All energy storage elements that exist in the system exist outside of the control loop.
  – For purely mechanical systems, separation between sensor and actuator is at most a rigid link.

• **Non-Collocated Control System**
  – At least one storage element exists inside the control loop.
  – For purely mechanical systems, separating link between sensor and actuator is flexible.
• Physical Interpretation of Poles and Zeros
  – Complex Poles of a collocated control system and those of a non-collocated control system are identical.
  – Complex Poles represent the resonant frequencies associated with the energy storage characteristics of the entire system.
  – Complex Poles, which are the natural frequencies of the system, are independent of the locations of sensors and actuators.
  – At a frequency of a complex pole, even if the system input is zero, there can be a nonzero output.
– Complex Poles represent the frequencies at which energy can freely transfer back and forth between the various internal energy storage elements of the system such that even in the absence of any external input, there can be nonzero output.

– Complex Poles correspond to the frequencies where the system behaves as an energy reservoir.

– Complex Zeros of the two control systems are quite different and they represent the resonant frequencies associated with the energy storage characteristics of a sub-portion of the system defined by artificial constraints imposed by the sensors and actuators.
– Complex Zeros correspond to the frequencies where the system behaves as an energy sink.

– Complex Zeros represent frequencies at which energy being applied by the input is completely trapped in the energy storage elements of a sub-portion of the original system such that no output can ever be detected at the point of measurement.

– Complex Zeros are the resonant frequencies of a subsystem constrained by the sensors and actuators.
Transfer Function Pole-Zero Example

Two-Mass, Three-Spring, Motor-Driven Dynamic System (shown with optical encoders instead of infrared position sensors)
Physical System Schematic

Two-Mass Three-Spring Dynamic System
Two-Mass Three-Spring Dynamic System
Physical Model
Mathematical Model:
Transfer Functions and State Space Equations

\[
\begin{align*}
\frac{X_1(s)}{V_{\text{in}}(s)} &= \frac{3.2503s^2 + 4.5887s + 2518.7}{s^4 + 2.3449s^3 + 1265.7s^2 + 1414.1s + 284460} \\
\frac{X_2(s)}{V_{\text{in}}(s)} &= \frac{1259.3}{s^4 + 2.3449s^3 + 1265.7s^2 + 1414.1s + 284460}
\end{align*}
\]

\[A = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
-489.45 & -0.93313 & 244.72 & 0 \\
0 & 0 & 0 & 0 & 1 \\
387.45 & 0 & -774.90 & -1.4118
\end{bmatrix} \quad B = \begin{bmatrix}
0 \\
3.2503 \\
0 \\
0
\end{bmatrix} \quad C = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix} \quad D = \begin{bmatrix}
0 \\
0
\end{bmatrix}\]
Mathematical Model: Transfer Functions and State Space Equations

Poles:

\[-0.536 \pm 17.1i \quad \Rightarrow \quad \omega = 17.1 \text{ rad} / \text{s} \quad \zeta = 0.0313\]
\[-0.637 \pm 31.2i \quad \Rightarrow \quad \omega = 31.2 \text{ rad} / \text{s} \quad \zeta = 0.0204\]

\[
\frac{X_1(s)}{V_{in}(s)} \Rightarrow -0.706 \pm 27.8i \quad \Rightarrow \quad \omega = 27.8 \text{ rad} / \text{s} \quad \zeta = 0.0254
\]

Zeros:

\[
\frac{X_2(s)}{V_{in}(s)} \Rightarrow \quad \text{None}
\]
Frequency Response Plots: Analytical

\[
\frac{X_1}{V_{in}}
\]

**Bode Diagrams**

- Frequency (rad/sec)
- Phase (deg); Magnitude (dB)

Graph showing the frequency response with peaks and valleys in the magnitude plot and a smooth phase plot.
Frequency Response Plots: Analytical

\[ \frac{X_2}{V_{in}} \]
Sensitivity Analysis

- Consider the function $y = f(x)$. If the parameter $x$ changes by an amount $\Delta x$, then $y$ changes by the amount $\Delta y$. If $\Delta x$ is small, $\Delta y$ can be estimated from the slope $\frac{dy}{dx}$ as follows:

$$\Delta y = \frac{dy}{dx} \Delta x$$

- The relative or percent change in $y$ is $\frac{\Delta y}{y}$. It is related to the relative change in $x$ as follows:

$$\frac{\Delta y}{y} = \frac{dy}{dx} \frac{\Delta x}{y} = \left( \frac{x}{y} \frac{dy}{dx} \right) \frac{\Delta x}{x}$$
• The sensitivity of $y$ with respect to changes in $x$ is given by:

$$S_x^y = \frac{x \ dy}{y \ dx} = \frac{d y / y}{d x / x} = \frac{d(\ln y)}{d(\ln x)}$$

• Thus

$$\frac{\Delta y}{y} = S_x^y \frac{\Delta x}{x}$$

• Usually the sensitivity is not constant. For example, the function $y = \sin(x)$ has the sensitivity function:

$$S_x^y = \frac{x \ dy}{y \ dx} = \frac{x \cos(x)}{y} = \frac{x \cos(x)}{\sin(x)} = \frac{x}{\tan(x)}$$
• **Sensitivity of Control Systems to Parameter Variation and Parameter Uncertainty**

  - A process, represented by the transfer function $G(s)$, is subject to a changing environment, aging, ignorance of the exact values of the process parameters, and other natural factors that affect a control process.

  - In the open-loop system, all these errors and changes result in a changing and inaccurate output.

  - However, a closed-loop system senses the change in the output due to the process changes and attempts to correct the output.

  - The sensitivity of a control system to parameter variations is of prime importance.
A primary advantage of a closed-loop feedback control system is its ability to reduce the system’s sensitivity.

Consider the closed-loop system shown. Let the disturbance \( D(s) = 0 \).

An open-loop system’s block diagram is given by:
The system sensitivity is defined as the ratio of the percentage in the system transfer function $T(s)$ to the percentage change in the process transfer function $G(s)$ for a small incremental change:

$$T(s) = \frac{C(s)}{R(s)}$$

$$S_G^T = \frac{\partial T / T}{\partial G / G} = \frac{\partial T}{\partial G} \frac{G}{T} = G_c(s) \frac{G(s)}{G_c(s)G(s)} = 1$$

For the open-loop system:

$$T(s) = \frac{C(s)}{R(s)} = G_c(s)G(s)$$

$$S_G^T = \frac{\partial T / T}{\partial G / G} = \frac{\partial T}{\partial G} \frac{G}{T} = G_c(s) \frac{G(s)}{G_c(s)G(s)} = 1$$
- For the closed-loop system

\[
T(s) = \frac{C(s)}{R(s)} = \frac{G_c(s)G(s)}{1 + G_c(s)G(s)H(s)}
\]

\[
S^T_G = \frac{\partial T / T}{\partial G / G} = \frac{\partial T}{\partial G} \frac{G}{T}
\]

\[
= \frac{1}{(1 + G_cGH)^2} \frac{G}{G_cG} = \frac{1}{G_c (1 + G_cGH)}
\]

- The sensitivity of the system may be reduced below that of the open-loop system by increasing \(G_cGH(s)\) over the frequency range of interest.
- The sensitivity of the closed-loop system to changes in the feedback element $H(s)$ is:

$$T(s) = \frac{C(s)}{R(s)} = \frac{G_c(s)G(s)}{1 + G_c(s)G(s)H(s)}$$

$$S_H^T = \frac{\partial T / T}{\partial H / H} = \frac{\partial T}{\partial H} \frac{H}{T}$$

$$= \frac{-(G_cG)^2}{(1 + G_cGH)^2} \frac{H}{G_cG} = \frac{-G_cGH}{1 + G_cGH}$$

- When $G_cGH$ is large, the sensitivity approaches unity and the changes in $H(s)$ directly affect the output response.
- Very often we seek to determine the sensitivity of the closed-loop system to changes in a parameter $\alpha$ within the transfer function of the system $G(s)$. Using the chain rule we find:

$$S_T^\alpha = S_G^T S_G^\alpha$$

- Very often the transfer function $T(s)$ is a fraction of the form:

$$T(s, \alpha) = \frac{N(s, \alpha)}{D(s, \alpha)}$$

- Then the sensitivity to $\alpha$ is given by:

$$S_T^\alpha = \frac{\partial T / T}{\partial \alpha / \alpha} = \frac{\partial \ln T}{\partial \ln \alpha} = \frac{\partial \ln N}{\partial \ln \alpha} \bigg|_{\alpha_0} - \frac{\partial \ln D}{\partial \ln \alpha} \bigg|_{\alpha_0} = S_N^\alpha - S_D^\alpha$$