MECH0010 Topic Notes

UCL

HD

October 28, 2020

Contents

Part I

Control

Chapter 1

Introduction to Control Systems

A **control system** is a mechanism which alters the future state of a particular system. **Control theory** is the method of selecting appropriate inputs. Control theory usually concerns dynamic behaviour of physical systems - the goal is to design a controller which leads to the system exhibiting the desired behaviour. Control systems have a large range of applications throughout engineering such as autopilot systems for ships and aircraft, radar tracking, robotics, machinery plant control and machine tools.

Methodology

- **Modelling** - Mathematical model of a system or "transfer function." Comes from detailed analysis of a system and often involves simplifications

- **Prediction** - The model is used to predict the behaviour of the system for a range of parameters and expected excitations

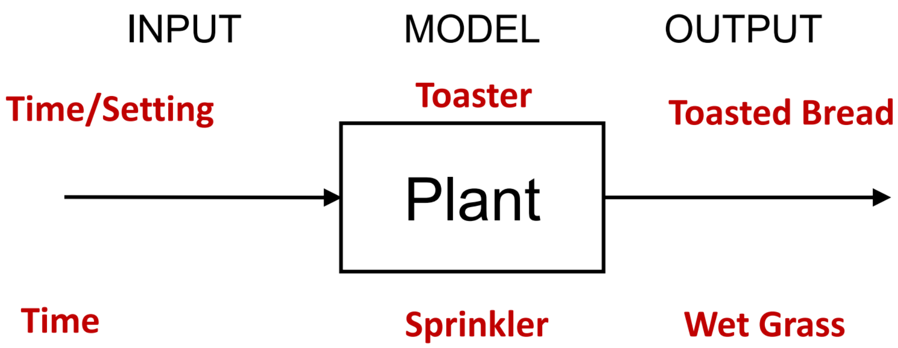
- **Design** - Design a controller which achieves its operating objectives and test it on the model

- **Test** - Here, theory is taken into reality and we compare our model to the physical hardware. Testing of the validity of our assumptions also takes place

- **Iterate** - Improve controller design through updated models

1.1 Open loop control

In an **open loop** control system, the input to the system is not dependant on any previous outputs. The output of the system is not being observed to confirm whether the desired output has been achieved.

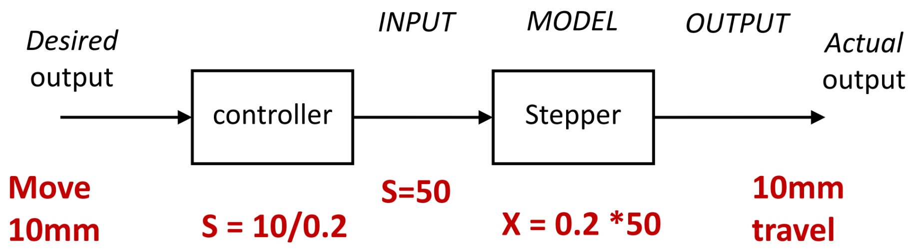


These control systems are simple and low cost. They are used in systems where feedback is not important. For example, a washing machine runs for 90 minutes, regardless of whether the clothes are dry after 60 minutes.

3D Printer example

Consider the stepper motors used to move the print head in x and y with belt and pulley system. Here, open loop control is used because the stepper motors are simple to control and have a relationship between input and output. For example, consider a stepper motor which moves 0.2mm each step (S) for an integer number of steps.

```
Plant model: X = 0.2 S
Control strategy: S = X/0.2
```



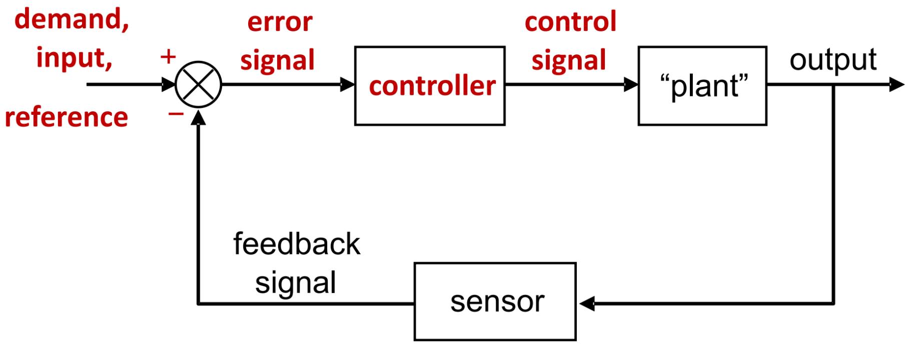
Implementing such a controller could be done with the following pseudocode:

```
Function MovePrintHeadX(Input_mm)
//convert to number of steps needed
Steps_S = Input_mm / 0.2
//send step command to motor
Stepper.step(Steps_S)
```

However, the assumption used to model our system is only valid if the stepper motors is within its specification i.e. friction, load, temperature, power are all nominal. As stated before, the system does not observe if the steps were successful or not.

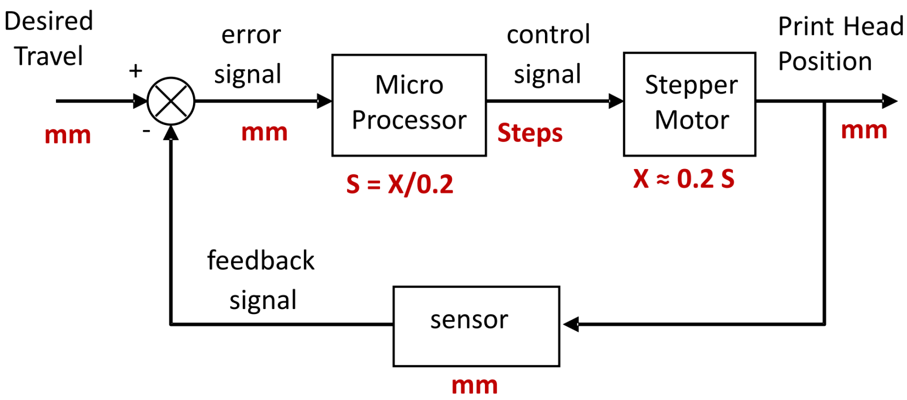
1.2 Closed loop control

By observing the output and comparing it against your desired output, it is possible to update the input to the system. The observed output is called the **feedback signal**.



The difference between the desired output and the feedback signal is known as the **error signal**. This is the signal that the controller uses. Feedback signals allow powerful controllers to be designed.

Taking our 3D printer for example, we can add a position sensor on the x-axis.

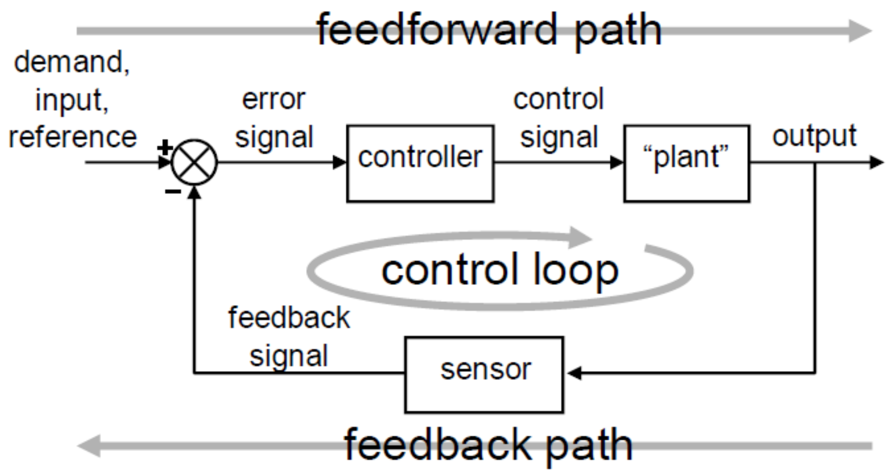


The pseudocode for this system could look like this

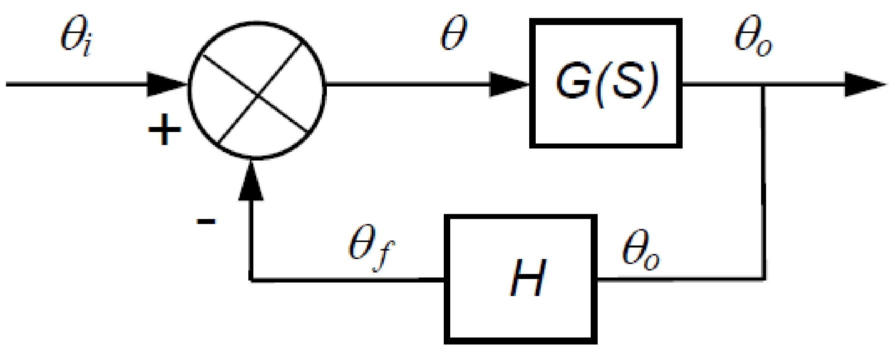
```
while
  // get measurement of position
  measured_value_mm = Sensor.getValue
  // calc error signal
  error_mm = setpoint_mm - measure_value_mm
  // convert to steps
  Steps_S - error_mm / 0.2
  // send step command to motor
  Stepper.Step(Steps_S)
end
```

We must consider whether this extra complexity is required or necessary in our system. Closed loop systems are not a magic bullet; they require careful modelling to predict system behaviour and a considered choice of parameters to prevent the system becoming unstable.

1.3 Block diagram representation of control loops



Control systems are often represented by block diagrams which show the information flow.



Flow path indicates the direction of data flow (from left to right). Values are maintained at a branch.

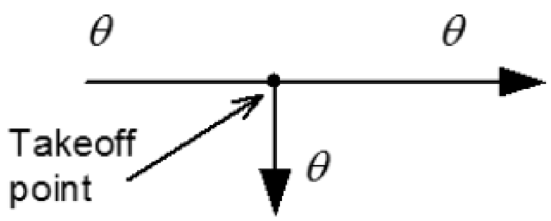


Figure 1.1: Flow path

Function block. The functions acts on the input to to produce the output. $x_{out} = f(x_{in})$

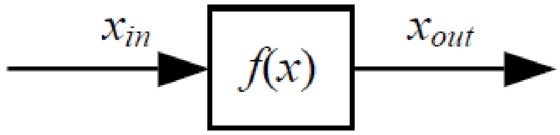


Figure 1.2: Function block

Comparator. The signals θ_1 and θ_2 are compared according to the signs (+ or -) and the result is θ_3. They **must** have the same units. In this case $+\theta_1 - \theta_2 = \theta_3$.

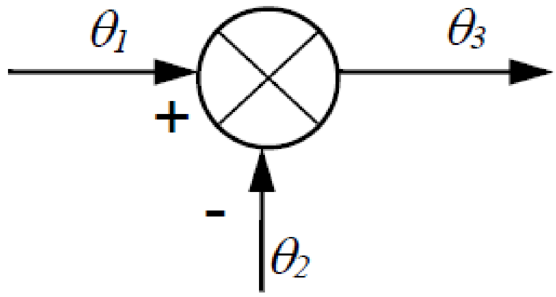
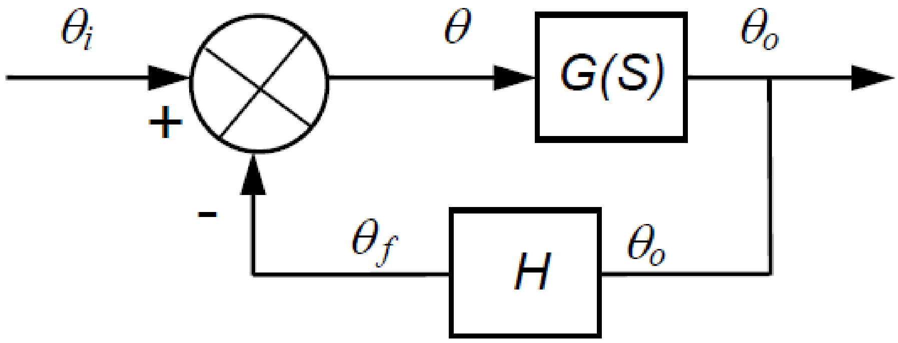


Figure 1.3: Comparator



1.4 Basic definitions/concepts

The following definitions are based on the standards of the IEEE (Institute of Electrical and Electronics Engineers)

- A **system** is an arrangement, set or collection of components connected or related in such a manner as to form an entirety or whole.
- A **control system** is an arrangement of physical components connected or related in such a manner as to command, direct or regulate itself or another

system. The components act together to perform a function not possible with any of the individual parts.

- The **input** is the stimulus or excitation applied to a system usually in order to produce a specified response.
- The **output** is the actual response obtained.
- An **open loop control system** is one in which the input is independent of the output.
- A **closed loop control system** is one in which the input is somehow dependent on the output.
- **Error signal** (or **actuating signal**) is the difference between the reference input and the feedback, in closed loop control systems it is this signal which is sent to the plant not the reference signal.

1.5 Notation

Large complex systems can be broken down into interconnected smaller ones and reduced further into a number of blocks.

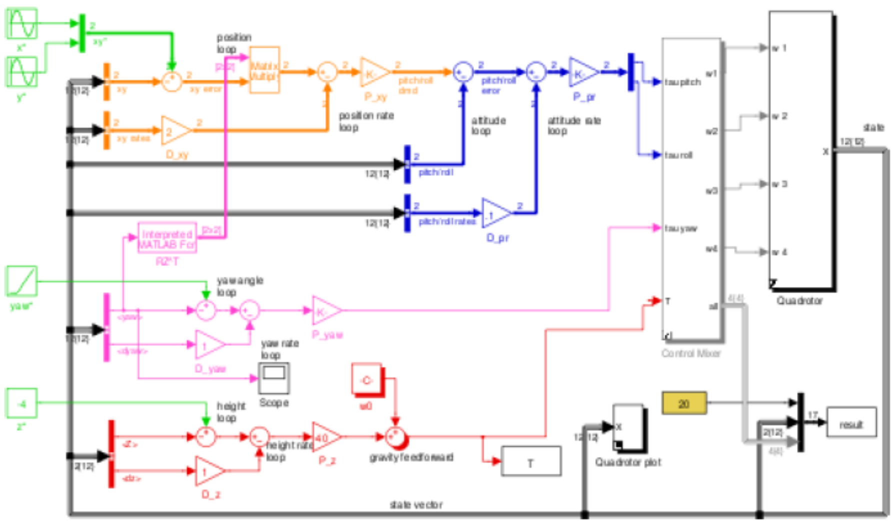
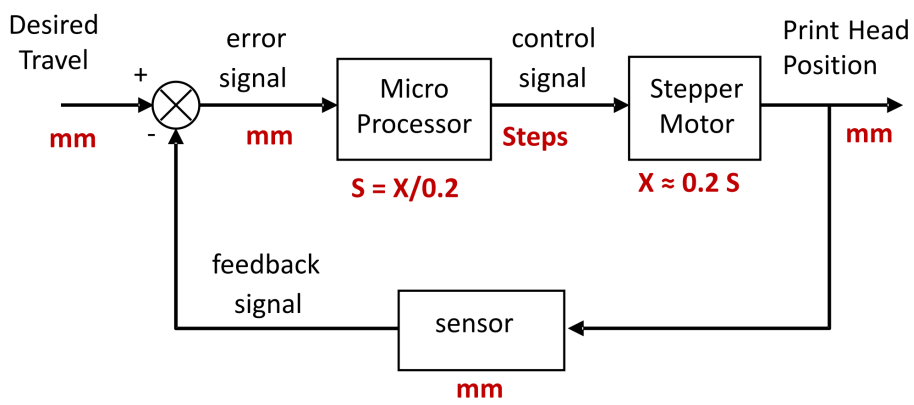
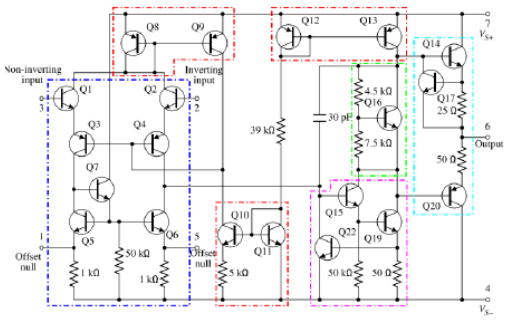
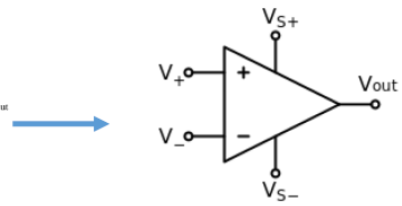


Figure 1.4: A quadcopter with lots of PID controllers running in parallel. Each colour represents a control loop for position, altitude, yaw etc.

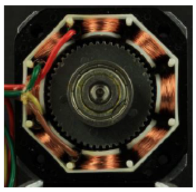
Complex systems can be abstracted to a signal block if the behaviour can be adequately modelled. Consider modelling the stepper motor from the 3D printer, in reality a complex device, as a single block. This is analogous to representing a complex op-amp circuit as a single unit.

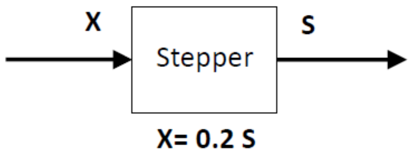






$$V_{\text{out}} = A_{OL}\,(V_+ - V_-)$$





1.6 Control system design process

1. Establish control goals
2. Identify the variables to control
3. Write the specifications for the variables
4. Establish the system configuration and identify the actuators
5. Obtain a model of the process, the actuator and the sensor
6. Describe a controller and select the parameters to be adjusted
7. Optimise the parameters and analyse the performance

1.7 Summary

- Control systems are interconnected components which are configured to provide a desired response.

- Two broad categories of control systems: open loop (no feedback of output) and closed loop (feedback signal).
- Successful design of the controller requires consideration of the design goals, definition of the specifications, system definition, modelling and subsequent analysis. Controller design is an iterative process.

1.8 Tutorial

3D printer design choice

The positioning of the print head is in open loop. Considering the comparative advantages of open and closed loop control, what would have lead the designers to make this choice?

Open loop control is used where the observation of the output is not necessary/relevant to the control of the system. Considering the accuracies required for 3D printing, which is around ± 0.05mm, we need to make sure that our print head moves to distances within this tolerance, otherwise our prints will be warped and misshaped. We must also consider that in an open system, small errors may lead to large errors over a certain period of operation, despite the print head moving distances within tolerance. There are also external factors affecting the belt and pulley system of the print head e.g. heat, dust, friction, load. These must all be within specification for the system to work properly.

Stepper motors are simple electronic components and can be manufactured to have a relationship between input and output. Hence, the need to observe whether the print head has moved a certain distance may be inconsequential. In applications where an extremely accurate print is required, a closed loop system would allow for a self calibrating system, which could allow the printer to work in different operational environments without the need to recalibrate or change components. However, this adds complexity to the controller and requires additional components to be added to the controller, increasing cost.

How might the designers overcome the problems mentioned without closed loop control?

Maintenance of the components in the system is important to make sure that everything is in the correct working order. When everything is kept clean and within the specification that the system is designed to work in, the controller will be able to work optimally. For example, calibrating the "0" position of the print head after cleaning the belts and pulleys from dust.

The designers may also implement a calibration operation into the controller before each phase. This may include implementing a "0" point into the hardware of the rails that the print head is attached to. The controller would move the print head to this position simply by moving the maximum distance left and down. Using this as a reference for each print would eliminate systematic distance errors.

Washing machine

Think of the control loops inside are they open or closed loop? Why? What could the sensors be?

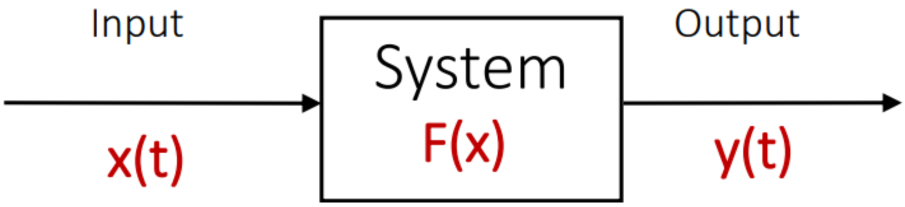
Let us consider the possible subsystems inside a washing machine.

- Timer - controls how long the washing machine operates. Open loop system. We only need a simple timing circuit with a minimum accuracy of around $\pm$1min. Even if the washing machine runs for 85 minutes instead of 90 (some degree of inaccuracy), this is insignificant.
- Motor - this is to spin the tub. Open loop system. Motors are simple electronic components with a relationship between input and output. An error signal provides no useful information for the operation of the motor.
- Temperature control - controls how hot or cold the water in the washing machine is. Open loop system. Most washing machines simply have a hot water connection for hot water and cold water. A solenoid valve can control the proportion of hot water to cold water that enters the tub. For washing machines with a heating element, the amount of power put through the coil will determine the final temperature of the water (relationship between input and output is known). Adding a thermistor to measure the output water temperature may be possible but ultimately unnecessary as the tolerance for the temperature can be quite high (± 5°C)
- Water valve control - controls how much water enters and leaves the drum. Closed loop control. The level of water in the drum needs to only be measured when the drum is full of water. Knowing the level of water at any given moment is unnecessary, thus using a simple pressure switch, will be sufficient in letting the controller know that the tub has reached capacity and close the input valve. After a set amount of time, the drain valve may open and since the amount of water in the tub when it is full is known, the drain time can simply be calculated as a function of its volume. This means that we do not need to check whether the water has drained from the tub.

Chapter 2

Modelling Linear Systems

Previously, we have considered systems in the general sense, with a generic function block approach.



Now we will investigate how we model physical systems and obtain the function block $F(x)$ for electrical and mechanical components.

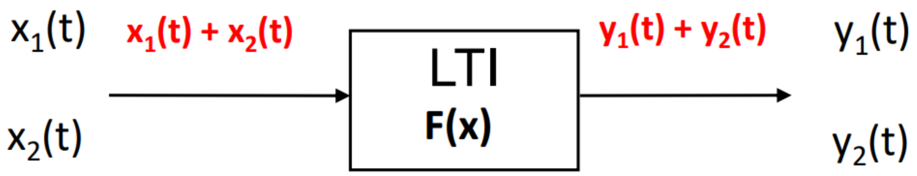
2.1 Linear Term Invariant (LTI) systems

The focus of this course and indeed much of control theory itself focuses on modelling physical systems as **linear** and **time invariant**. LTI systems have three key properties:

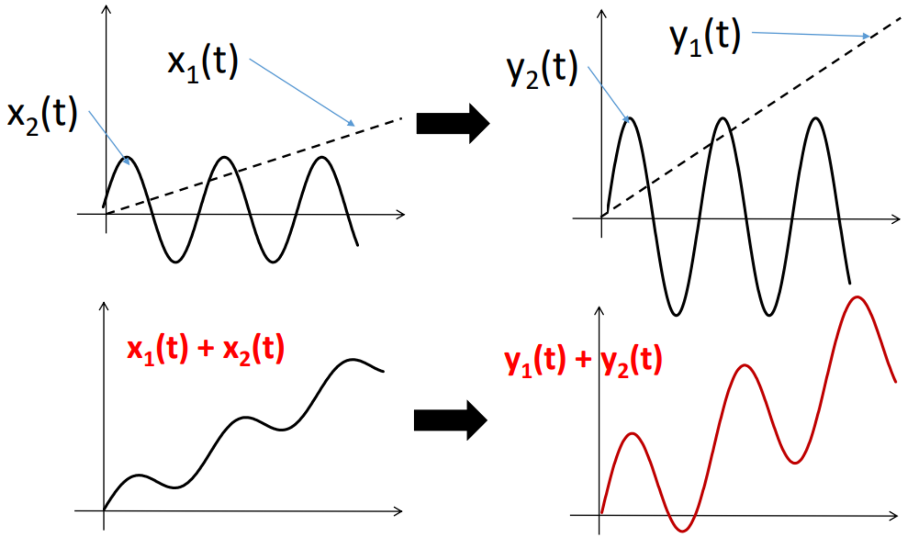
- Obey principle of superposition
- Homogeneity
- Time invariance

2.1.1 Superposition

If input $x_1(t)$ produces output $y_1(t)$ and input $x_2(t)$ produces $y_2(t)$, then input $x_1(t) + x_2(t)$ produces output $y_1(t) + y_2(t)$

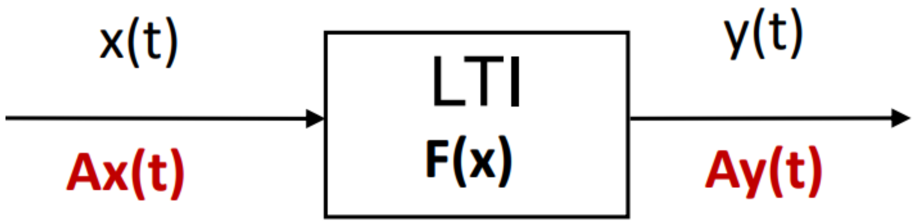


Say for a system which doubles the input $F(x) = 2x$

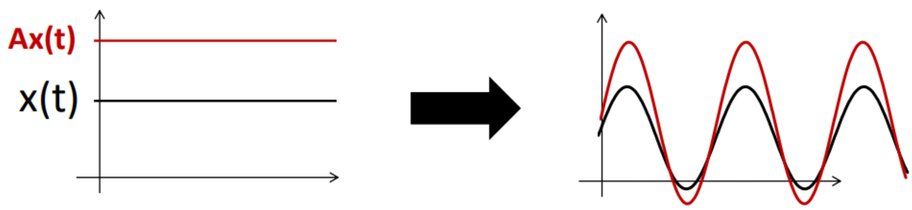


2.1.2 Homogeneity

If the input to the system $x(t)$ is scaled by a magnitude scale factor A, then the output $y(t)$ is also scaled by the same factor.

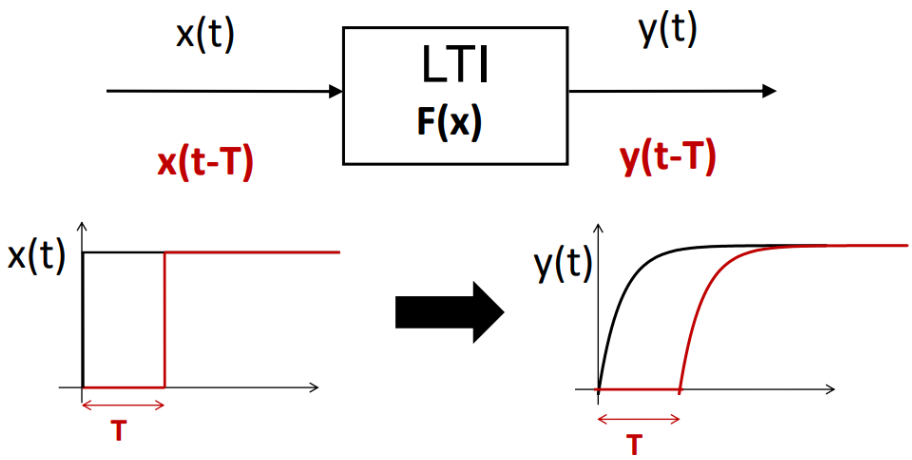


For example, consider a system which generates a sine wave at a given amplitude, with a set frequency:



2.1.3 Time invariance

If input is applied at time $t = 0$ or T s from now, the output is identical with the exception of a delay of T s.



Are these models suitable for physical systems?

These three requirements, whilst simple, are so stringent that **almost no physical LTI system truly exists**. Consider a car engine - the performance deteriorates over time, to stretch it further, would you expect a system to give the same output after a time delay T of 10 years? Even simple systems such as a resistor in an electrical circuit have non-linearities - a scaling factor A could be chosen for $x(t)$ which would mean too much current flows and the resistor melts.

Most practical systems are not linear, but often we can assume they behave linearly **under certain conditions/assumptions**. Linear systems are **much** easier to solve! There are **analytic** solutions with standard tools used solve the equations. Whereas for non-linear problems it is often necessary to solve them numerically.

2.1.4 Linearisation: Example 1

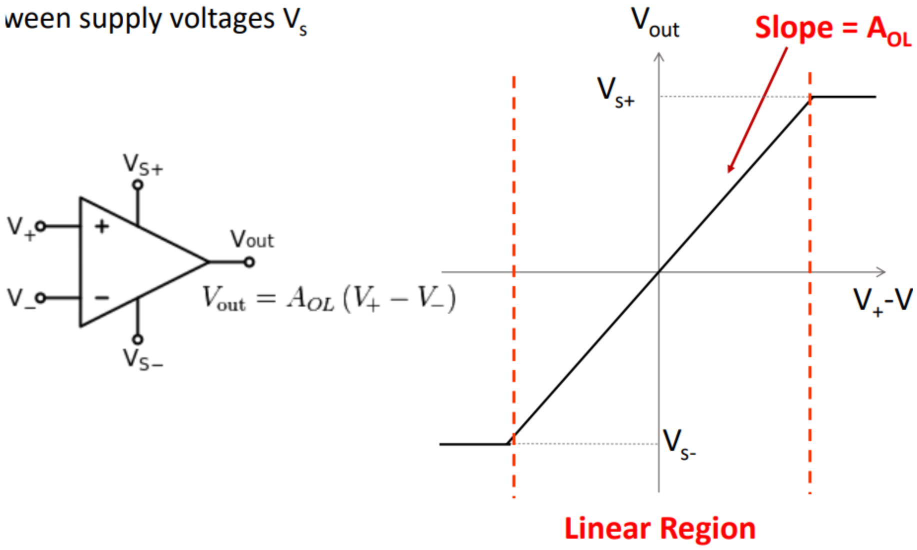
For a simple system such as a spring, across all possible compressions or extensions the response is non-linear:

$$F = -kX \tag{2.1}$$

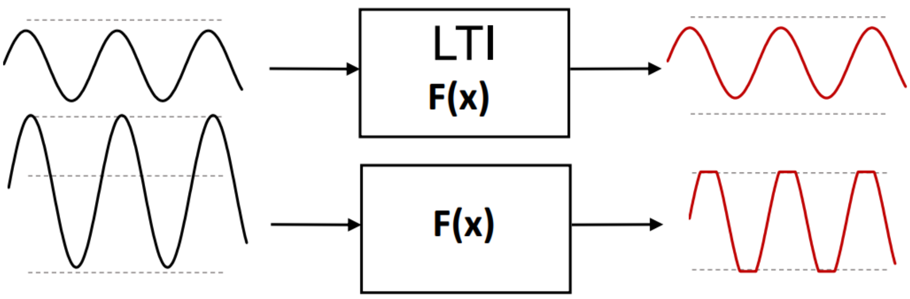
Hookes law is only a linear approximation of the true response. However, if we choose the operating range of the spring correctly, the response is within the linear region. This approximation is **valid**.

Linearisation: Example 2

Similarly an op-amp has a **linear region** where the output signals are between supple voltages V_S



Keep signals within these regions and the linear assumption holds:

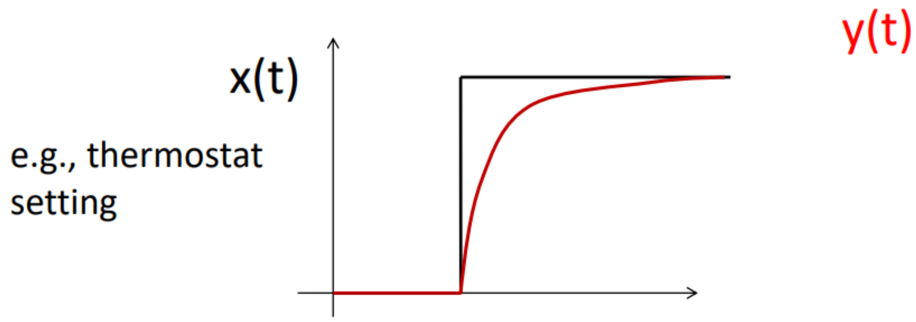


Add LTI example

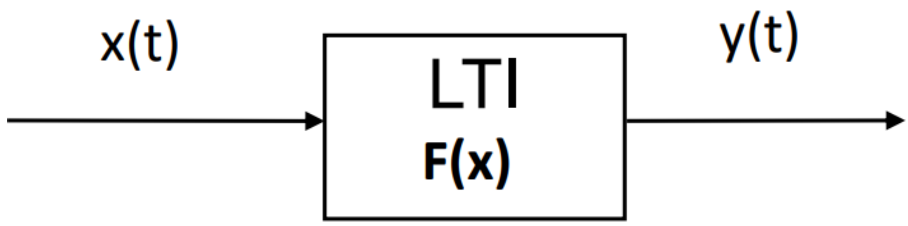
2.2 Dynamic systems - Laplace Transform

2.2.1 Dynamic systems as ODEs

Ideal systems would respond **instantaneously** to inputs, however real world systems require some time to adjust to changes and are thus known as **dynamic systems** as the output changes over time.



As we are intersected in describing something that **changes** with time, it is useful to express the function block of the system $F(t)$ as an ordinary differential equation (ODE).



$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + ... + a_2\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + a_1\frac{\mathrm{d}y}{\mathrm{d}t} + a_0 = bx \tag{2.2}$$

- x is input function or forcing function
- y is output
- n is **order** of the ODE
- a_0... are coefficients. These **completely characterise the system**

2.2.2 Laplace Transforms

Because of our **linear assumptions** we can use Laplace transforms to simplify solving the ODEs. The Laplace transforms of a signal (function) x is the function $X = \mathcal{L}(x)$ defined by

$$X(s) = \int_{0^-}^{\infty} x(t)e^{-st}\,\mathrm{d}t \tag{2.3}$$

For those $s \in \mathbf{C}$ for which the integral makes sense.

- X is a complex-valued function of complex numbers
- s is called the (complex) **frequency variable** with units s^{-1}, t is called the **time variable** (in sec); st is unitless

- $s = \sigma + j\omega$

As we shall see:

- Differential operators are replaced with algebraic variables
- Algebraic equations are much easier to manipulate & solve
- Standard forms exist for many physical systems

Laplace Transforms: Example 1

Let's fine Laplace transform $x(t) = e^t$:

$$X(e^t) = \int_{0^-}^{\infty} e^t e^{-st} \, \mathrm{d}t \tag{2.4}$$

$$X(e^t) = \int_{0^-}^{\infty} e^{(1-s)t} \, \mathrm{d}t \tag{2.5}$$

$$Xe^t = \frac{1}{1-s} e^{(1-s)t} \Big|_{0^-}^{\infty} \tag{2.6}$$

$$X(e^t) = \frac{1}{1-s} \times 0 - \frac{1}{1-s} \times 1 = \frac{1}{s-1} \tag{2.7}$$

Laplace Transforms: Example 2

Constant or **unit step** $x(t) = 1$ (for $t \geq 0$)

$$X(s) = \int_{0^-}^{\infty} 1e^{-st} \, \mathrm{d}t \tag{2.8}$$

$$X(s) = \int = -\frac{1}{s} e^{-st} \Big|_{0^-}^{\infty} \tag{2.9}$$

$$X(s) = -\frac{1}{s} \times 0 - (-\frac{1}{s}) \times 1 = \frac{1}{s} \tag{2.10}$$

Laplace Transforms: Example 3

Sinusoid: first express $x(t) = \cos(\omega t)$ as:

$$x(t) = \frac{1}{2} e^{i\omega t} + \frac{1}{2} e^{-i\omega t} \text{ (Euler's formula)} \tag{2.11}$$

$$X(s) = \int_{0^-}^{\infty} e^{-st} \left(\frac{1}{2} e^{i\omega t} + \frac{1}{2} e^{-i\omega t} \right) \mathrm{d}t \tag{2.12}$$

$$X(s) = \frac{1}{2} \int_{0^-}^{\infty} e^{(-s+i\omega)t} \, \mathrm{d}t + \frac{1}{2} \int_{0^-}^{\infty} e^{(-s-i\omega)t} \, \mathrm{d}t \tag{2.13}$$

$$X(s) = \frac{1}{2} \frac{1}{s - i\omega} + \frac{1}{2} \frac{1}{s + i\omega} = \frac{s}{s^s + \omega^s} \tag{2.14}$$

You can look up these transforms in a table. The Laplace variable, s, can be considered to represent the differential operator (very useful for control engineering):

$$s \equiv \frac{\mathrm{d}}{\mathrm{d}t} \tag{2.15}$$

$$\frac{1}{s} \equiv \int_{0^-}^{\infty} \mathrm{d}t \tag{2.16}$$

Laplace transform of time derivative $\frac{\mathrm{d}x}{\mathrm{d}t}$:

$$L\left\{\frac{\mathrm{d}x}{\mathrm{d}t}\right\} = \int_{0^-}^{\infty} \frac{\mathrm{d}x}{\mathrm{d}t} e^{-st}\,\mathrm{d}t \tag{2.17}$$

Integrating by parts

$$L\left\{\frac{\mathrm{d}x}{\mathrm{d}t}\right\} = s\int_{0^-}^{\infty} x(t)e^{-st}\,\mathrm{d}t + \left[x(t)e^{-st}\right]_{0^-}^{\infty} \tag{2.18}$$

The initial condition $x(0^-)$ is often zero in practice

$$L\left\{\frac{\mathrm{d}x}{\mathrm{d}t}\right\} = sX(s) - x(0^-) = sX(s) \tag{2.19}$$

We can substitute this result to solve higher order derivatives:

$$L\left\{\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}\right\} = sL\left\{\frac{\mathrm{d}x}{\mathrm{d}t}\right\} - \frac{difx}{\mathrm{d}t}(0^-) \tag{2.20}$$

$$L\left\{\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}\right\} = s^2X(s) - sx(0^-) - \frac{\mathrm{d}X}{\mathrm{d}t}(0^-) = s^2X(s) \tag{2.21}$$

So more generally, with all initial conditions set to zero:

$$L\left\{\frac{d^n x}{dt^n}\right\} = s^n X(s) \tag{2.22}$$

2.2.3 Transfer Functions

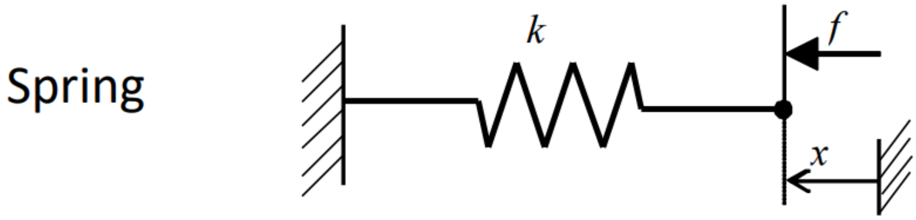
After we have taken the Laplace transform of the differential equation of a system, it's useful to rearrange to give the system **output** as the product of the system **input** and the system **transfer function**.

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

2.2.4 Transfer Functions of Mechanical Components

Spring

Convention is input force, output displacement. Balance **forces**



Time domain equation (Hooke's Law):

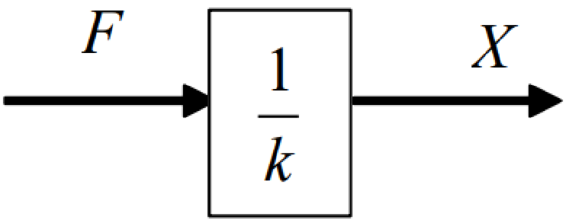
$$f(x) - kx \tag{2.23}$$

k is stiffness in $\mathrm{N\,m^{-1}}$. Laplace domain equation:

$$F(s) = kX(s) \tag{2.24}$$

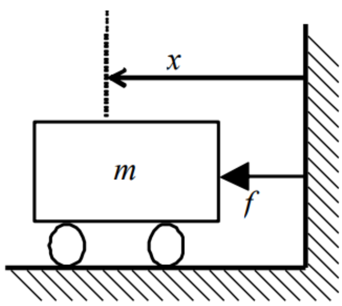
Transfer function:

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{k} \tag{2.25}$$



Mass

Inertial load - mass



Time domain equation from Newton's 2nd law of motion

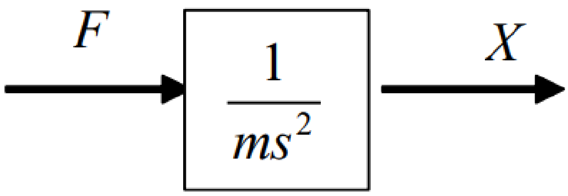
$$f(t) = m\frac{d^2 x(t)}{dt^2} \tag{2.26}$$

Laplace domain equation

$$F(s) = ms^2 X(s) \tag{2.27}$$

Transfer function

$$\frac{X(s)}{F(s)} = \frac{1}{ms^2} \tag{2.28}$$



Damper

Below is a dashpot - a viscous damper. They resist motion through friction. The damping coefficient is in terms of c with units $\mathrm{N\,m^{-1}\,s^{-1}}$. Time domain equation:

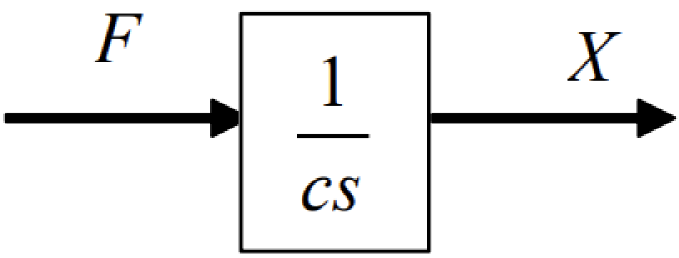
$$f(x) = x\frac{\mathrm{d}x}{\mathrm{d}t} = csx \tag{2.29}$$

Laplace domain equation:

$$F(s) = csX(s) \tag{2.30}$$

Transfer function

$$\frac{X(s)}{F(s)} = \frac{1}{cs} \tag{2.31}$$



2.2.5 Combining components

Springs

Springs in parallel

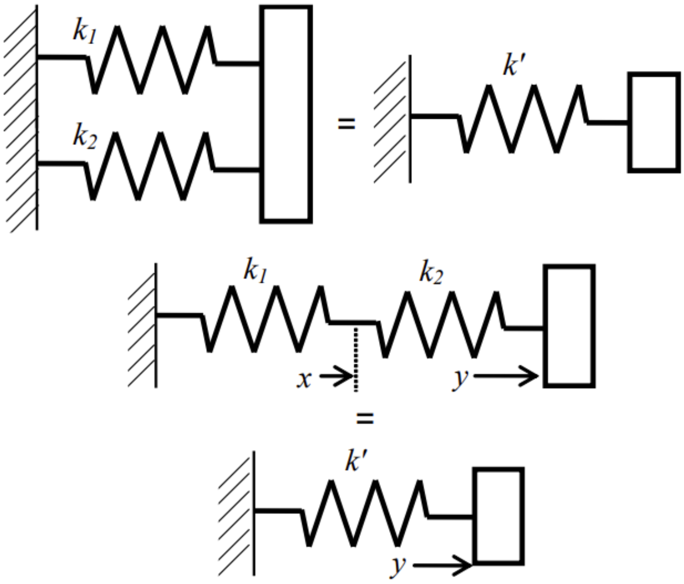
$$k' - k_1 + k_2 \tag{2.32}$$

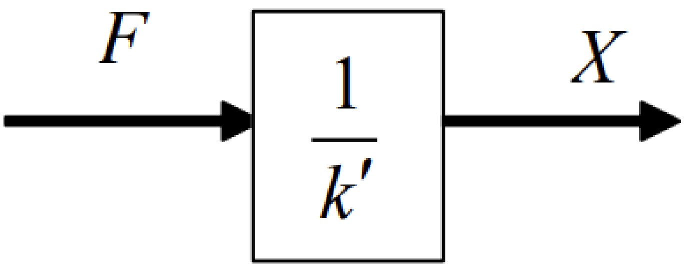
$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{k'} \tag{2.33}$$

Springs in series

$$\frac{1}{k'} = \frac{1}{k_1} + \frac{1}{k_2} \tag{2.34}$$

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{k'} \tag{2.35}$$





Dashpots behave in a similar way to springs. Parallel:

$$c' = c_1 + c_2 \tag{2.36}$$

and in series:

$$\frac{1}{c'} = \frac{1}{c_1} + \frac{1}{c_2} \tag{2.37}$$

Resistor

Convention is input voltage, output current. Balance voltages. Ohms law is $V = IR$. Time domain:

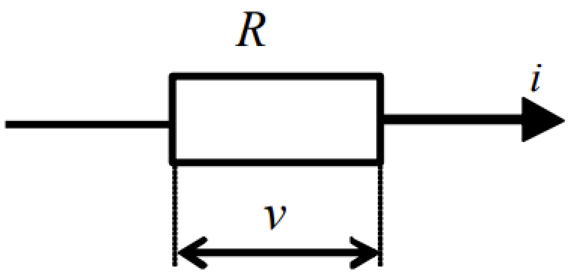
$$v(t) = i(t)R \tag{2.38}$$

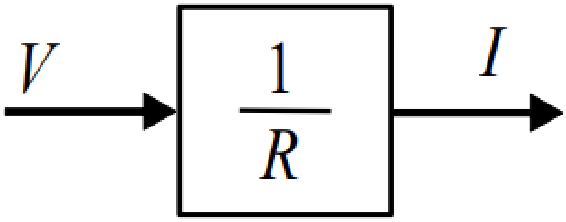
Laplace domain:

$$V(s) = I(s)R \tag{2.39}$$

Transfer function:

$$G(s) = \frac{I(s)}{V(s)} = \frac{1}{R} \tag{2.40}$$





Capacitor

Either definition of current/voltage relationship gives same result. Time domain:

$$i(t) = C\frac{\mathrm{d}v}{\mathrm{d}t} \tag{2.41}$$

$$v(t) = \frac{1}{C}\int i(t)\,\mathrm{d}t \tag{2.42}$$

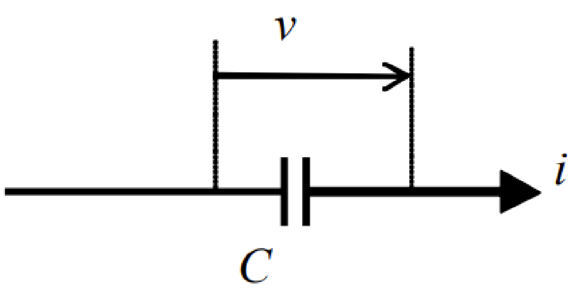
Laplace domain:

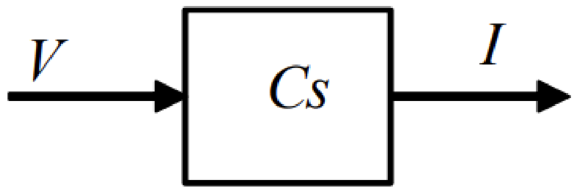
$$I(s) = CsV(s) \tag{2.43}$$

$$V(s) = \frac{1}{C}\frac{1}{s}I(s) \tag{2.44}$$

Transfer function:

$$G(s) = \frac{I(s)}{V(s)} = Cs \tag{2.45}$$





Inductor

An inductor resists changes of current by generating a voltage in opposition via magnetic induction. From Faraday's Law:

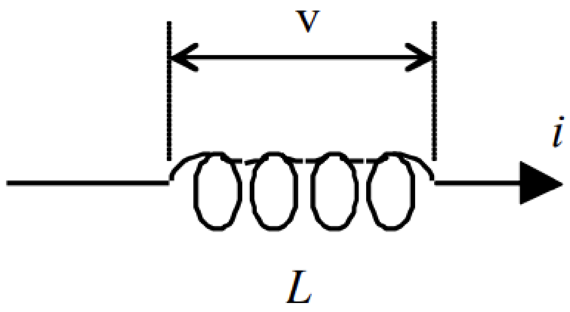
$$v(t) = L\frac{\mathrm{d}i}{\mathrm{d}t} \tag{2.46}$$

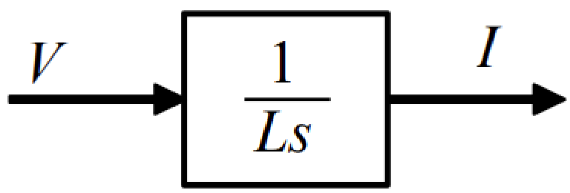
Laplace domain

$$V(s) = LsI(s) \tag{2.47}$$

Transfer function:

$$G(s) = \frac{I(s)}{V(s)} = \frac{1}{Ls} \tag{2.48}$$





Chapter 3

First and Second Order Systems

19/10/2020

3.1 First Order Systems

All first order systems i.e. those with only $\frac{\mathrm{d}x}{\mathrm{d}t}$ take the following "standard" forms:

$$\frac{X(s)}{Y(s)} = \frac{\alpha}{1+Ts} = \frac{\gamma}{1+\tau s} \tag{3.1}$$

Where:

- τ, γ is the **gain**
- T, τ is the **time constant**

This function is commonly known as an exponential time delay, or lag.

3.1.1 Returning to the Time Domain

To get the response of the system in the time domain we need to convert the transfer function from the Laplace domain. Tables of common inverse Laplace functions have already been created; all we need to do arrange our transfer function into one of these standard forms.

Laplace transforms − Table			
$f(t) = L^{-1}\{F(s)\}$	$F(s)$	$f(t) = L^{-1}\{F(s)\}$	$F(s)$
$a \quad t \geq 0$	$\frac{a}{s} \quad s > 0$	$\sin \omega t$	$\frac{\omega}{s^2+\omega^2}$
$at \quad t \geq 0$	$\frac{a}{s^2}$	$\cos \omega t$	$\frac{s}{s^2+\omega^2}$
e^{-at}	$\frac{1}{s+a}$	$\sin(\omega t + \theta)$	$\frac{s \sin\theta + \omega \cos\theta}{s^2+\omega^2}$
te^{-at}	$\frac{1}{(s+a)^2}$	$\cos(\omega t + \theta)$	$\frac{s \cos\theta - \omega \sin\theta}{s^2+\omega^2}$
$\frac{1}{2}t^2e^{-at}$	$\frac{1}{(s+a)^3}$	$t \sin \omega t$	$\frac{2\omega s}{(s^2+\omega^2)^2}$
$\frac{1}{(n-1)!}t^{n-1}e^{-at}$	$\frac{1}{(s+a)^n}$	$t \cos \omega t$	$\frac{s^2-\omega^2}{(s^2+\omega^2)^2}$
e^{at}	$\frac{1}{s-a} \quad s > a$	$\sinh \omega t$	$\frac{\omega}{s^2-\omega^2} \quad s > \lvert\omega\rvert$
te^{at}	$\frac{1}{(s-a)^2}$	$\cosh \omega t$	$\frac{s}{s^2-\omega^2} \quad s > \lvert\omega\rvert$
$\frac{1}{b-a}(e^{-at}-e^{-bt})$	$\frac{1}{(s+a)(s+b)}$	$e^{-at}\sin \omega t$	$\frac{\omega}{(s+a)^2+\omega^2}$
$\frac{1}{a^2}[1-e^{-at}(1+at)]$	$\frac{1}{s(s+a)^2}$	$e^{-at}\cos \omega t$	$\frac{s+a}{(s+a)^2+\omega^2}$
t^n	$\frac{n!}{s^{n+1}} \quad n = 1,2,3..$	$e^{at}\sin \omega t$	$\frac{\omega}{(s-a)^2+\omega^2}$
t^ne^{at}	$\frac{n!}{(s-a)^{n+1}} \quad s > a$	$e^{at}\cos \omega t$	$\frac{s-a}{(s-a)^2+\omega^2}$
t^ne^{-at}	$\frac{n!}{(s+a)^{n+1}} \quad s > a$	$1-e^{-at}$	$\frac{a}{s(s+a)}$
$\sqrt{t}$	$\frac{\sqrt{\pi}}{2s^{3/2}}$	$\frac{1}{a^2}(at-1+e^{-at})$	$\frac{1}{s^2(s+a)}$
$\frac{1}{\sqrt{t}}$	$\sqrt{\frac{\pi}{s}} \quad s > 0$	$f(t-t_1)$	$e^{-t_1 s}F(s)$
$g(t) \cdot p(t)$	$G(s) \cdot P(s)$	$f_1(t) \pm f_2(t)$	$F_1(s) \pm F_2(s)$
$\int f(t)dt$	$\frac{F(s)}{s} + \frac{f^{-1}(0)}{s}$	$\delta(t)$ *unit impulse*	$1 \quad$ *all s*
$\frac{df}{dt}$	$sF(s) - f(0)$	$\frac{d^2f}{df^2}$	$s^2F(s) - sf(0) - f'(0)$
$\frac{d^nf}{dt^n}$	$s^nF(s) - s^{n-1}f(0) - s^{n-2}f'(0) - s^{n-3}f''(0) - \cdots - f^{n-1}(0)$		

Figure 3.1: Laplace Transformation Table

Example: Parallel Spring & Damper (Laplace Domain found previously)

$$\frac{X(s)}{Y(s)} = \frac{\alpha}{1+Ts} = \frac{\gamma}{1+\tau s} \tag{3.2}$$

Looking at the table, we can find the relevant inverse function; in this case (3rd from the top):

$$a = \frac{1}{\tau} \tag{3.3}$$

Hence:

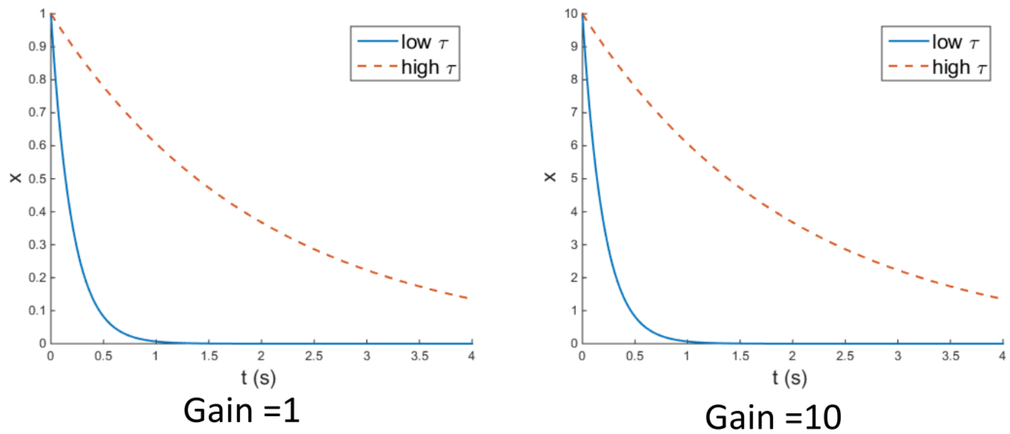
$$\frac{X(s)}{Y(s)} = \frac{\frac{\gamma}{\tau}}{\frac{1}{\tau}+s} = \frac{\gamma}{\tau}\frac{1}{\frac{1}{\tau}+s} \tag{3.4}$$

$$x(t) = L^{-1}\left\{\frac{\gamma}{\tau}\frac{1}{\frac{1}{\tau}+s}\right\} \tag{3.5}$$

$$= \frac{\gamma}{\tau}L^{-1}\left\{\frac{1}{\frac{1}{\tau}+s}\right\} \tag{3.6}$$

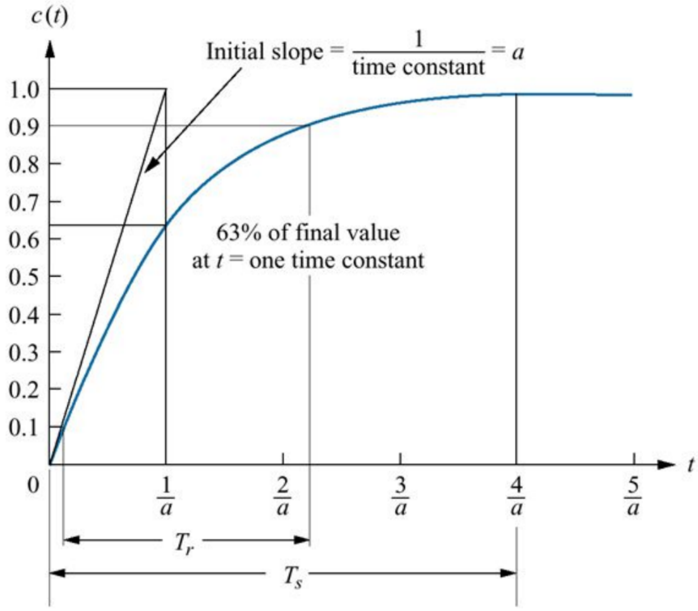
$$= \frac{\gamma}{\tau}e^{-\frac{t}{\tau}} \tag{3.7}$$

This is an exponential decay – the only variable needed to define the system is the time constant. The gain merely scales the response.



Having a negative time constant makes no physical sense, so the exception of $T = 0$, the response will always be an exponential decay.

3.1.2 Unit Step Response



If a step input is given to a first order system, this will result in an exponential increase. This might be something like changing the room temperature. When talking about unit step input, the amplitude of the input will vary from 0 to 1. Some important definitions are:

- Initial slope $= \frac{1}{time\ constant} = a$
- T_r = Raising time = how quick the exponential function achieves 90% of the final value
- τ = Time constant = how quick the exponential function achieves 63% of the final value
- T_s = Settling time = how quick the exponential function becomes constant

Remember, we have modelled mechanical and electrical systems, and arrived at the same equations. This could be biological, financial, meteorological... This is good, because you only need to understand one system to understand hundreds. However, this means the models can appear abstract (or possibly arbitrary), so it is important to keep in mind the physical systems we are trying to model.

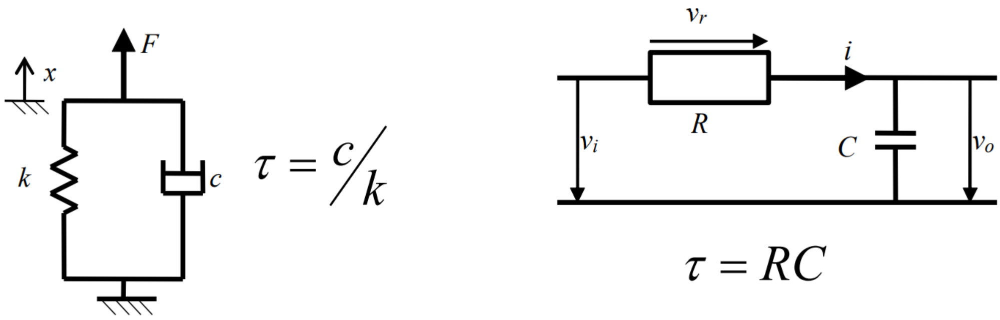


Figure 3.2: Left: Rate of decay determined by ratio of spring return force to viscous friction — Right: Rate of decay determined by total impedance of R and C

3.2 Second Order Systems

The standard form for second order systems is shown below:

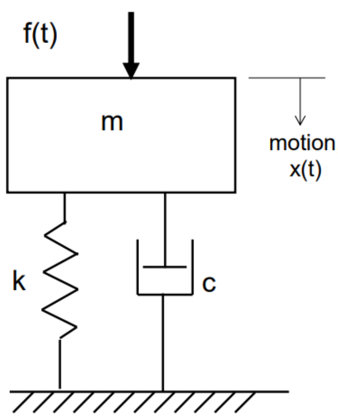
$$G(s) = \gamma \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{3.8}$$

Where:

- γ is the **gain**
- ω_n is the **natural frequency**
- ζ is the **damping ratio**

This function is known as a damped oscillator, in that it produces harmonic sinusoidal oscillations which decay over time. This type of system appears everywhere in physics, as well as in engineering. Even to the extent that some higher order systems are simplified to become second order, just because it is so well understood.

Example – Mass Spring Damper



Consider a simple mechanical system Mass/Spring/Damper (MSD). As before, we wish to relate the input force to the output displacement i.e. Transfer function desired:

$$G(s) = \frac{X(s)}{F(s)} \tag{3.9}$$

Balancing forces as function of time:

$$f(t) = f_S(t) + f_D(t) + f_I(t) \tag{3.10}$$

Where:

- $f_S(t) = kx(t)$
- $f_D(t) = c\frac{\mathrm{d}x(t)}{\mathrm{d}t}$
- $f_I(t) = m\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2}$

So time domain force equilibrium is:

$$f(t) = kx(t) + c\frac{\mathrm{d}x(t)}{\mathrm{d}t} + m\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} \tag{3.11}$$

Converting to Laplace domain:

$$F(s) = kX(s) + csX(s) + ms^2X(s) \tag{3.12}$$

$$F(s) = X(s)(k + cs + ms^2) \tag{3.13}$$

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{ms^2 + cs + k} \tag{3.14}$$

To obtain the Laplace domain in the "standard" form, the coefficient of the highest order of s on the denominator should be 1:

$$G(s) = \frac{\frac{k}{m}}{s^2 + s\frac{c}{m} + \frac{k}{m}}\frac{1}{k} \tag{3.15}$$

$$= \gamma\frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{3.16}$$

Where:

- $\omega_n = \sqrt{\frac{k}{m}}$
- $\zeta = \frac{c}{2\sqrt{km}}$
- $\gamma = \frac{1}{k}$

3.2.1 Returning to the Time Domain

As before, we use the inverse Laplace transform to get the time domain response. To reiterate, the benefit of standard forms is that the transforms are given in the tables:

$$G(s) = \gamma \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{3.17}$$

$$x(t) = L^{-1}\left\{\frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}\right\} \tag{3.18}$$

$$\frac{x(t)}{x(0)} = \gamma \frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} sin(\omega_n\sqrt{1-\zeta^2}\cdot t) \tag{3.19}$$

γ is not shown as it is unaffected by the Laplace transform. The equation above looks complicated, but investigating each term yields:

$$\gamma \frac{\omega_n}{\sqrt{1-\zeta^2}} \tag{3.20}$$

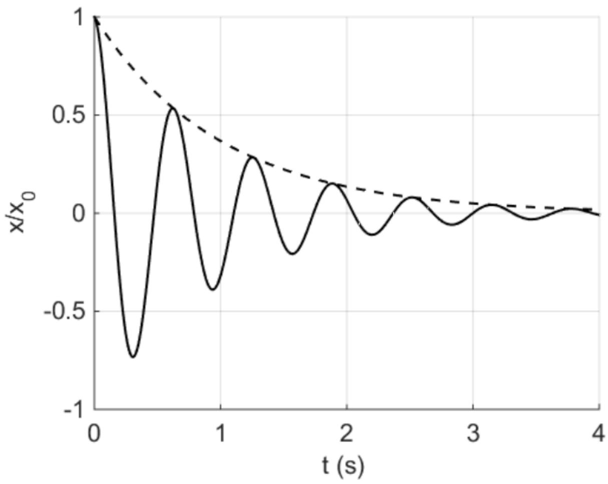
γ ω_n ζ are all constants, so the whole term is just a number.

$$e^{-\zeta\omega_n t} \tag{3.21}$$

It is an exponential function, which depending on whether $-\zeta\omega_n$ is positive or negative, increases or decays.

$$sin(\omega_n\sqrt{1-\zeta^2}\cdot t) \tag{3.22}$$

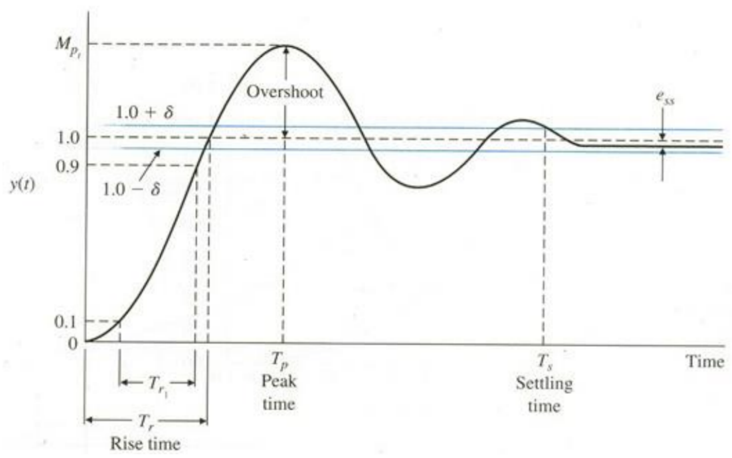
ω_n ζ are just constants, so assuming $\zeta < 1$, this is just a sine wave.



This is an exponentially decaying sinusoidal oscillation, with:

- Frequency: $\omega_n\sqrt{1-\zeta^2}\cdot t$
- Decay: $e^{-\zeta\omega_n t}$
- Gain: $\gamma\frac{\omega_n}{\sqrt{1-\zeta^2}}$

3.2.2 Unit Step Response

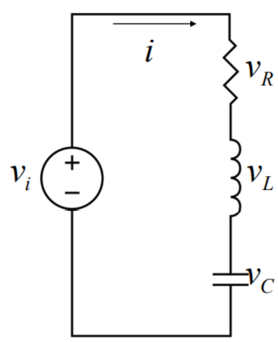


Some properties that should be known, but will be investigated later are:

- Slope at the time step input is made $(t = 0) = 0$
- T_p = Peak Time
- T_s = Settling Time
- T_r = Rise Time
- Overshoot

3.3 Tutorial

LRC Filter



A more complex filtering circuit is an LRC filter, also known as a harmonic oscillator circuit, which allows for a narrower range of frequencies to be amplified or attenuated. These circuits are used in analogue radios with a variable capacitor or inductor connected to the dial to select the frequency. To describe the relationship between the input voltage V_i, and the voltage across the capacitor V_C:

$$G(s) = \frac{V_C(s)}{V_i(s)} \tag{3.23}$$

Balancing voltages using Kirchoff's Law:

$$v_i = v_R + v_L + v_C \tag{3.24}$$

Where:

- $v_r(t) = i(t)R$
- $v_c(t) = \frac{1}{C}\int i(t)\,\mathrm{d}t$
- $v_L(t) = L\frac{\mathrm{d}i(t)}{\mathrm{d}t}$

So time domain voltage is:

$$v_i(t) = i(t)R + \frac{1}{C}\int i(t)\,\mathrm{d}t + L\frac{\mathrm{d}i(t)}{\mathrm{d}t} \tag{3.25}$$

Converting to Laplace domain:

$$V_i(s) = I(s)R + \frac{1}{Cs}I(s) + LsI(s) \tag{3.26}$$

$$V_i(s) = I(s)\left(R + \frac{1}{Cs} + Ls\right) \tag{3.27}$$

Getting the laplace domain and rearranging the $v_c(t)$ term yields:

$$v_c(t) = \frac{1}{C}\int i(t)\,\mathrm{d}t \tag{3.28}$$

$$V_c(s) = \frac{1}{Cs}I(s) \tag{3.29}$$

$$I(s) = V_C(s)Cs \tag{3.30}$$

Inputting equation (3.30) into equation (3.27):

$$V_i(s) = CsV_c(s)\left(R + \frac{1}{Cs} + Ls\right) \tag{3.31}$$

$$G(s) = \frac{V_C(s)}{V_i(s)} = \frac{1}{LCs^2 + CRs + 1} \tag{3.32}$$

To obtain the Laplace domain in the "standard form":

$$G(s) = \frac{\frac{1}{LC}}{s^2 + s\frac{R}{L} + \frac{1}{LC}} \tag{3.33}$$

$$= \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{3.34}$$

Where:

- $\omega_n = \sqrt{\frac{1}{LC}}$
- $\zeta = \frac{R}{2}\sqrt{\frac{C}{L}}$
- $\gamma = 1$

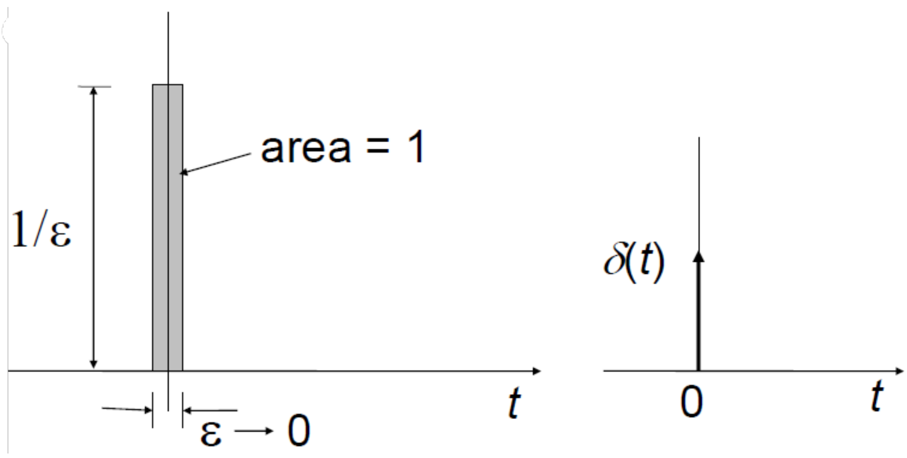
Chapter 4

System Time Response

4.1 Impulse functions/responses

4.1.1 Impulse response of a system: Dirac delta function

A useful tool in analysing the transient response of a system is the impulse signal, a unit (amplitude = 1) pulse infinitesimally small, with area = 1. Formally this is known as the **Dirac delta function (impulse function)**.



The Dirac delta function is a non-physical, singularity function with the follow definition:

$$\delta(t) = \begin{cases} 0 \text{ for } t \neq 0 \\ \text{undefined for } t = 0 \end{cases} \tag{4.1}$$

but with the requirement that

$$\int_{-\infty}^{\infty} \delta(t)\,\mathrm{d}t = 1 \tag{4.2}$$

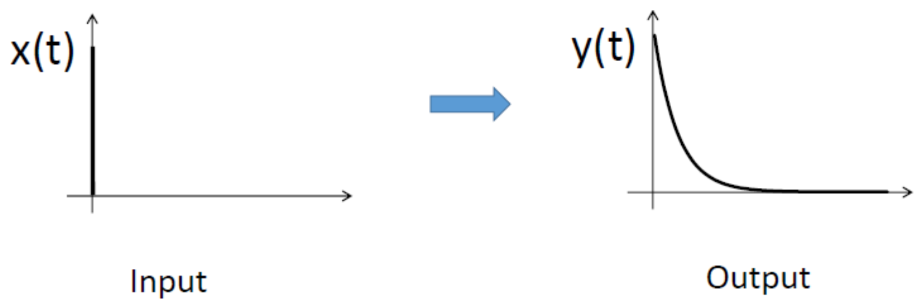
so taking the Laplace transform of this is also just 1

$$\mathcal{L}(\delta(t)) = \int_{0^-}^{\infty} \delta(t)e^{-st}\,\mathrm{d}t = 1 \tag{4.3}$$

Thus, the impulse response of the system is equal to the transfer function and from this it can be shown that **any** arbitrary signal can be described as a summation of impulse responses.

4.1.2 Impulse response on a first order system

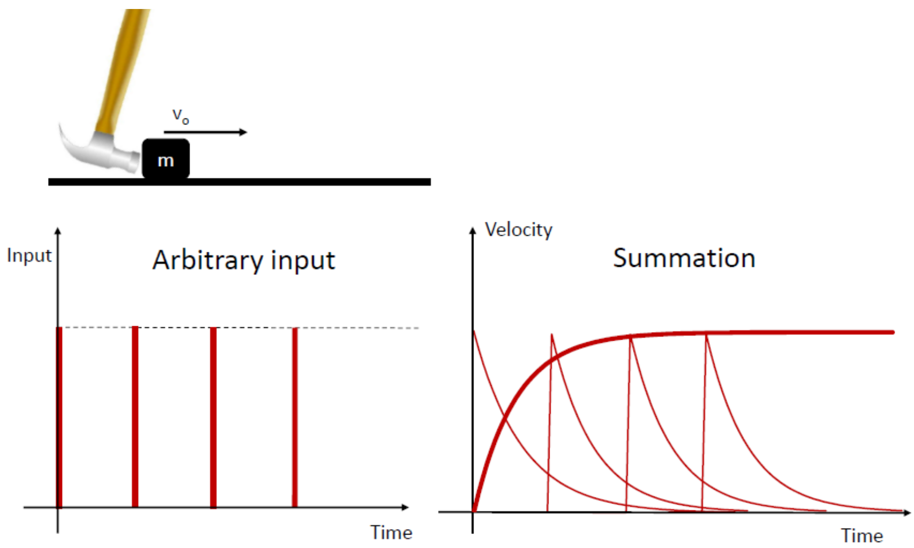
Take a first order response for example, the transfer function and thus the impulse response looks like this:



Due to our LTI assumptions:

- Scaling the input scales the output
- Superposition of inputs equals superposition of outputs
- Time invariance

4.1.3 Impulse response of a system



Time vs Frequency Domain

- u is the impulse function to the system
- h is called the impulse response of the system
- H is called the transfer function (TF) of the system

$$y(t) = \int_0^\infty h(\tau)u(t-\tau)\,\mathrm{d}\tau = \int_0^\infty h(\tau - t)u(\tau)\,\mathrm{d}\tau \tag{4.4}$$

with $0 \leq \tau \leq t$

$$y(t) = h(t) \cdot u(t) \tag{4.5}$$

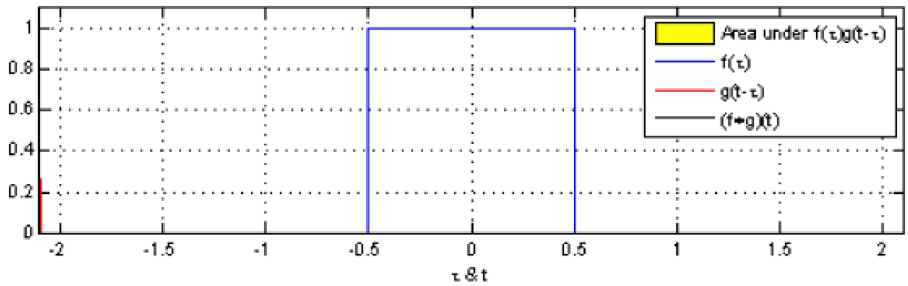
This is called convolution.

$$Y(s) = H(s) \cdot U(s) \tag{4.6}$$

This is called multiplication.

Convolution example

Essentially, the steps for convolving two signals are to first reflect the signal g, then offset the reflected signal. Then calculate the area under the graph for every offset, by sliding $-g$. The convolution at each time point is equal to the area under the intersection of functions. For two pulses, the result is a triangle wave:



However, the calculations to obtain this result in the time domain are complicated, but are only multiplication in the Laplace domain.

Getting the Time Response

The procedure to describe the time response for LTI systems is thus:

- Express the input, $u(t)$, in Laplace notation, $U(s)$
- Use this to find output $Y(s)$, usually by multiplying $U(s)$ by the transfer function $Y(s) = U(s)G(s)$
- Use inverse Laplace transforms (from tables) to express $Y(s)$ as a function of time, $y(t)$

4.2 Input functions: Impulse - Step - Ramp

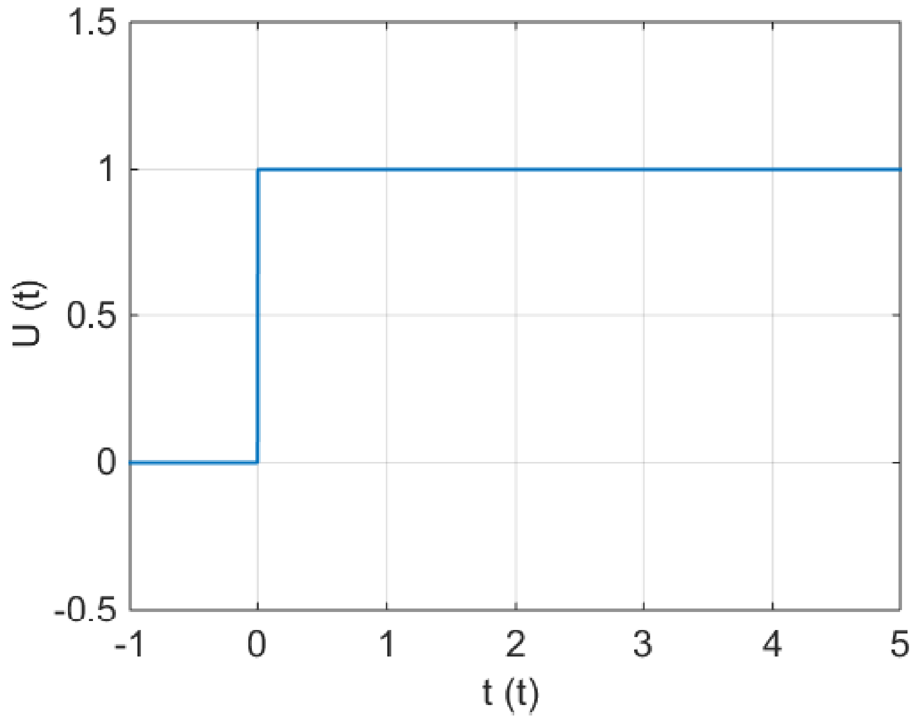
Now we will consider some standard inputs and look at the response of first and second order systems:

- Impulse
  - The Laplace transform is 1, so the response to an impulse is by the definition the transfer function
- Step
- Ramp

There are many others, particularly sinusoidal inputs or other discontinuous inputs, which are important in control loops, but we will focus on the two classic examples.

4.2.1 Step input

A step input is a discontinuous function, which is zero for all negative values of t and 1 for all positive values.



Laplace transform

$$x(t) = U(t) \tag{4.7}$$

$$\mathcal{L}\left\{U(t)\right\} = \int_{0-}^{\infty} e^{-st}\,\mathrm{d}t = \left[-\frac{1}{s}e^{-st}\right]_{0-}^{\infty} \to \frac{1}{s} \tag{4.8}$$

$$\tag{4.9}$$

Or for a gain of A

$$x(t) = AU(t) \tag{4.10}$$

$$\mathcal{L}\left\{AU(t)\right\} = \frac{A}{s} \tag{4.11}$$

Applications

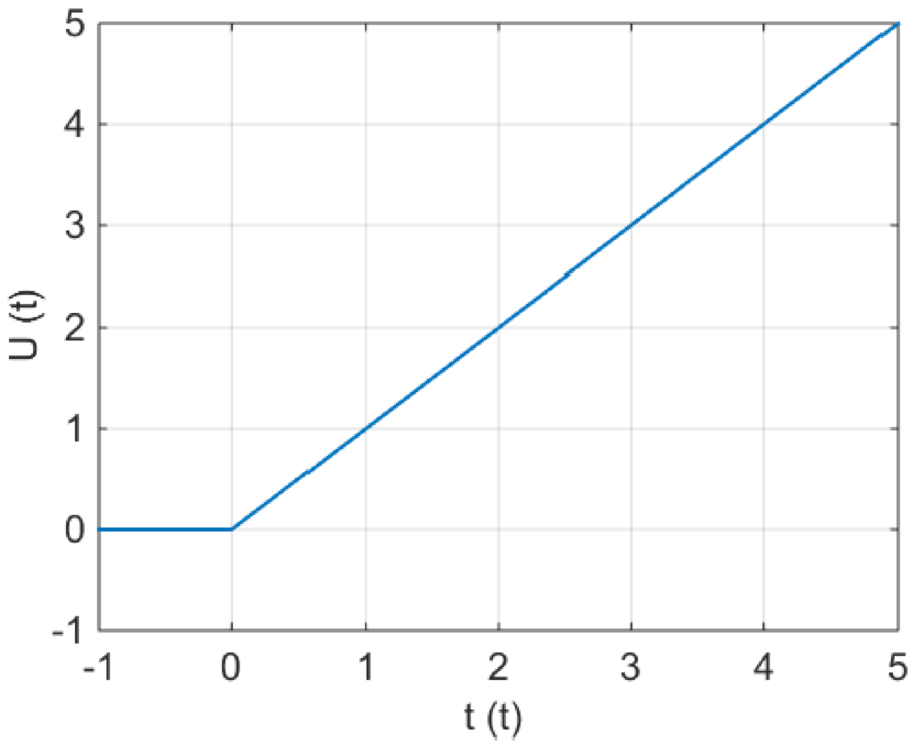
The step response is extremely useful in control theory for describing the behaviour of the system. In part because it incorporates the "transient" behaviour - from the sudden change from zero to one, as well as the "steady state" behaviour as the system settles down to a single value. IT also replicates many real world control applications such as:

- Position control - move to a $X = 10$mm position and stay
- Speed control - go to 33 RPM
- Temperature - heat element on 3D printer to 230 °C

Also, unlike the Dirac impulse - it is physically realisable.

4.2.2 Ramp input

A ramp input has a value of t for all t values above zero and zero elsewhere, often is scaled by a gain A.



Laplace transform

$$x(t) = at \tag{4.12}$$

$$\mathcal{L}\left\{at\right\} \int_{0^-}^{\infty} ate^{-st}\,\mathrm{d}t = -a\left[\frac{t}{s}e^{-st}\right]_0^{\infty} + a\int_0^{\infty}\frac{1}{s}e^{-st}\,\mathrm{d}t \tag{4.13}$$

$$= a\left[-\frac{1}{s^2}e^{-st}\right] \rightarrow \frac{a}{s^2} \tag{4.14}$$

Applications

Ramp inputs are useful in understanding the steady state behaviour of a system i.e. when t goes to infinity. Practical examples of control applications using ramp inputs are

- Servo motors - shaft **position** rather than speed

- Ovens for PCB manufacturing etc. - strict linear **profile** of temperature required as opposed to "get to this temperature quickly"
- CNC milling machine, move in X direction and constant rate

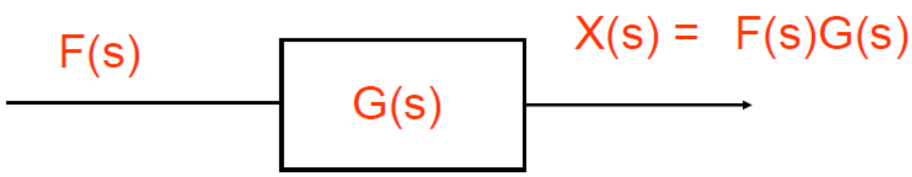
4.2.3 Summary of Input Functions

$$\text{Impulse } F(s) = A \tag{4.15}$$

$$\text{Step } F(s) = \frac{A}{s} \tag{4.16}$$

$$\text{Ramp } F(s) = \frac{A}{s^2} \tag{4.17}$$

For a **unit** response, $A = 1$. We can apply these inputs to the LTI system by multiplying the transfer function by the input, both in terms of s.



4.3 First Order System Step Response

Electromagnets are used in motors both rotary and linear, as well as in power transfer and also magnetic levitation in trains. First obtain the transfer function.

$$\frac{I(s)}{V(s)} \tag{4.18}$$

Balancing voltage gives:

$$v(t) = v_R(t) + v_I(t) \tag{4.19}$$

where

$$v_R(t) = i(t)R \tag{4.20}$$

$$v_I(t) = L\frac{\mathrm{d}i(t)}{\mathrm{d}t} \tag{4.21}$$

Substitution gives:

$$v(t) = Ri(t) + L\frac{\mathrm{d}i(t)}{\mathrm{d}t} \tag{4.22}$$

In the Laplace domain

$$V(s) = RI(s) + LsI(s) \tag{4.23}$$

$$\frac{I(s)}{V(s)} = \frac{1}{(Ls + R)} \tag{4.24}$$

For a **step** input of 1 V (unit step)

$$V(s) = \frac{1}{s} \tag{4.25}$$

$$I(s) = V(s)\frac{1}{(Ls+R)} \tag{4.26}$$

$$= \frac{1}{s(Ls+R)} \tag{4.27}$$

Sadly there is no direct equivalent in the tables for this transform. The first takes is to split the expression up using partial fractions, into expressions that are give in the tables.

$$I(s) = \frac{1}{s(Ls+R)} = \frac{\frac{1}{L}}{s(s+\frac{R}{L})} \text{ (remove coeff for s)} \tag{4.28}$$

$$= \frac{k_1}{s} + \frac{k_2}{(s+\frac{R}{L})} \text{ partial fraction expansion} \tag{4.29}$$

where

$$k_1(s+\frac{R}{L}) + k_2 s = \frac{1}{L} \tag{4.30}$$

Looking at $s = 0$

$$k_1\frac{R}{L} = \frac{1}{L} \rightarrow k_1 = \frac{1}{R} \tag{4.31}$$

Looking at $s = -\frac{R}{L}$

$$-k_2\frac{R}{L} = \frac{1}{L} \rightarrow k_2 = -\frac{1}{R} \tag{4.32}$$

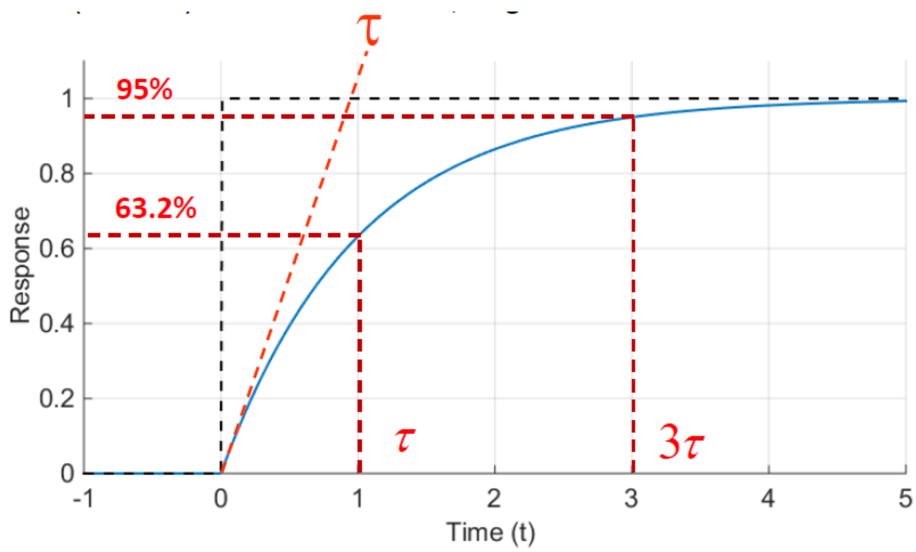
Which yields

$$I(s) = \frac{1}{R}\left[\frac{1}{s} - \frac{1}{(s+\frac{R}{L})}\right] \tag{4.33}$$

Looking at the Laplace tables, we can now use entries for $\frac{1}{s}$ and $\frac{1}{(s+a)}$, considering $t > 0 \therefore u(t) = 1$.

$$I(t) = \frac{1}{R}\left[1 - e^{-\frac{R}{L}t}\right] \tag{4.34}$$

Which has the familiar form of a first order exponential rise, with a steady state gain of $\frac{1}{R}$. Time constant $\tau = \frac{L}{R}$. For now assume that L and $R = 1$, so gain and time constant are 1.



Thus for a known time constant, it is possible to calculate the response at any point after the step. A common parameter of a first order system is the **settling time** which is the time taken to reach 95% of the desired value, or 3τ. The inverse process is also used in system identification, because the response is so simple - we can measure the step response and calculate the time constant of the system. If we do convolution in the time domain you get the same too.

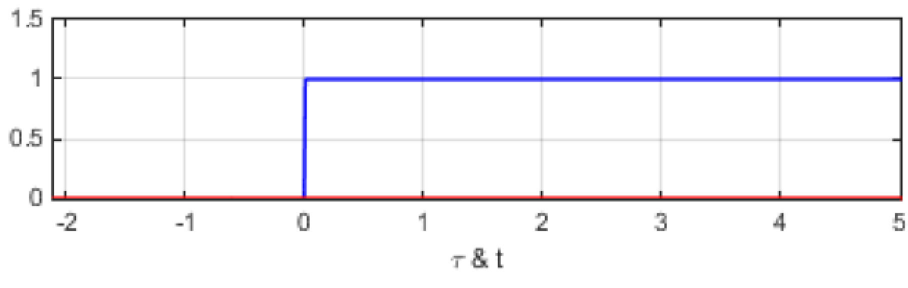
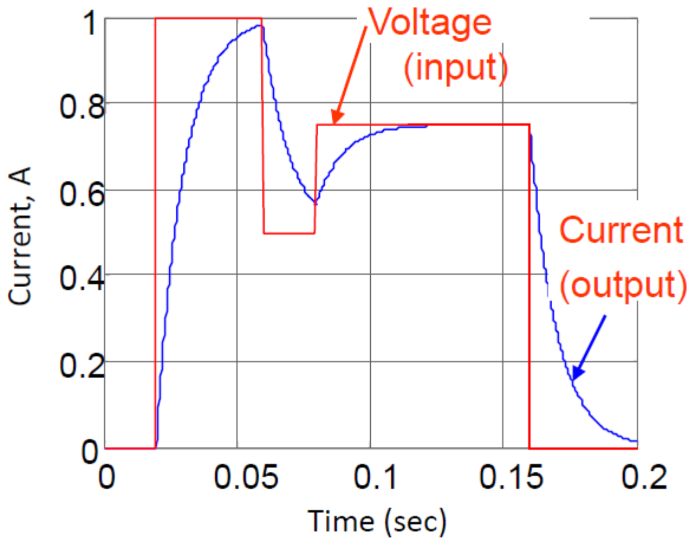


Figure 4.1: Area under $f(\tau)g(1-\tau)$, blue line $f(\tau)$, red line $g(t-\tau)$, black line $(f+g)t$.

These "step responses" can be repeated for different values of V and for different starting values of $I(t)$. The response will always be behind the input for all time constant > 0, and thus these systems are referred to as **first order lags**.



4.4 First Order System Ramp Response

For a **ramp** input of gain 1 (unit ramp)

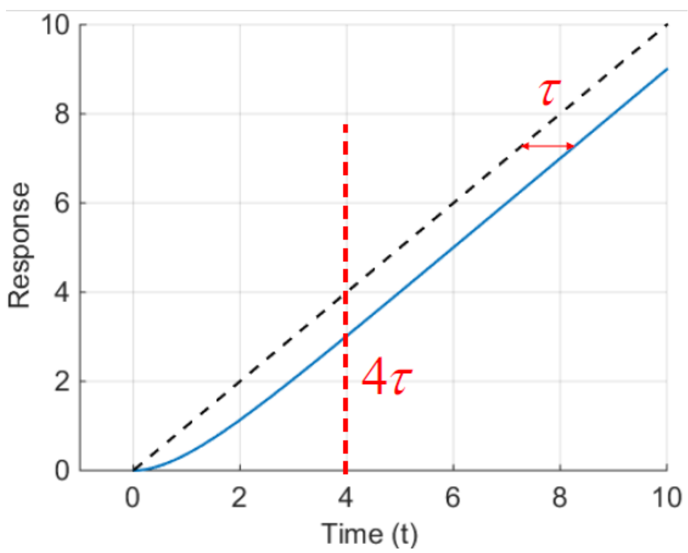
$$\frac{I(s)}{V(s)} = \frac{1}{(Ls+R)} \tag{4.35}$$

$$V(s) = \frac{1}{s^2} \tag{4.36}$$

$$I(s) = \frac{1}{s^2(Ls+R)} = \frac{\frac{1}{L}}{s^2(s+\frac{R}{L})} \tag{4.37}$$

Again, this is not in the Laplace tables, so rearrange use partial fractions, giving us the general solution:

$$I(t) = t - \tau\left(1 - e^{-\frac{t}{\tau}}\right) \tag{4.38}$$



After 4 time constants, we can say that the system is in the **steady state** where the difference between the input and output is no longer changing. The steady state lag is then equal to τ

4.5 Understanding Poles and Zeros

As defined, the transfer function is a rational function in the complex variable $s = \sigma + jw$, that is

$$G(s) = \frac{b_m s^m + b_{m-1}s^{m-1} + ... + b_1 s + b_0}{a_n s^n + a_{n-1}s^{n-1} + ... + a_1 s + a_0} \tag{4.39}$$

It is often convenient to factor the polynomials in the numerator and denominator, and to write the transfer function terms of those factors:

$$G(s) = \frac{N(s)}{D(s)} = K\frac{(s - z_1)(s - z_2)...(s - z_{m-1})(s - z_m)}{(s - p_1)(s - p_2)...(s - p_{n-1})(s - p_n)} \tag{4.40}$$

Where the numerator and denominator polynomials, $N(s)$ and $D(s)$, have real coefficients defined by the system's differential equation and $K = \frac{b_m}{a_n}$. Poles and zeros are found by

$$N(s) = 0 \text{ and } D(s) = 0 \tag{4.41}$$

All of the coefficients of polynomials $N(s)$ and $D(s)$ are real, therefore the poles and zeros must be either purely real, or appear in complex conjugate pairs.

> The poles and zeros are properties of the transfer function and therefore of the differential equation describing the input-output system dynamics. Together with the gain constant, they completely characterize the differential equation, and provide a complete description of the system.

4.5.1 Example

A linear system is described by the differential equation:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 2\frac{\mathrm{d}y}{\mathrm{d}t} + 5y = 3\frac{\mathrm{d}u}{\mathrm{d}t} + 12u \tag{4.42}$$

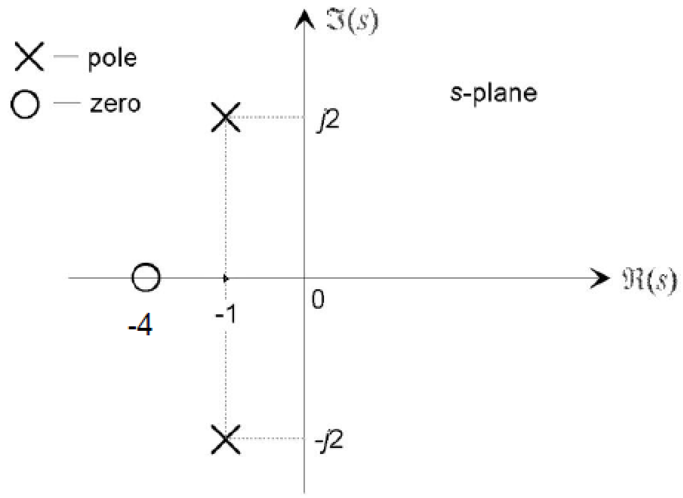
$$\mathcal{L}\left\{\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 2\frac{\mathrm{d}y}{\mathrm{d}t} + 5y\right\} = s^2Y(s) + 2sY(s) + 5Y(s) \tag{4.43}$$

$$\mathcal{L}\left\{3\frac{\mathrm{d}u}{\mathrm{d}t} + 12u\right\} = 3sU(s) + 12U(s) \tag{4.44}$$

$$N(s) = 3s + 12 = 0 \rightarrow z_1 = -4 \tag{4.45}$$

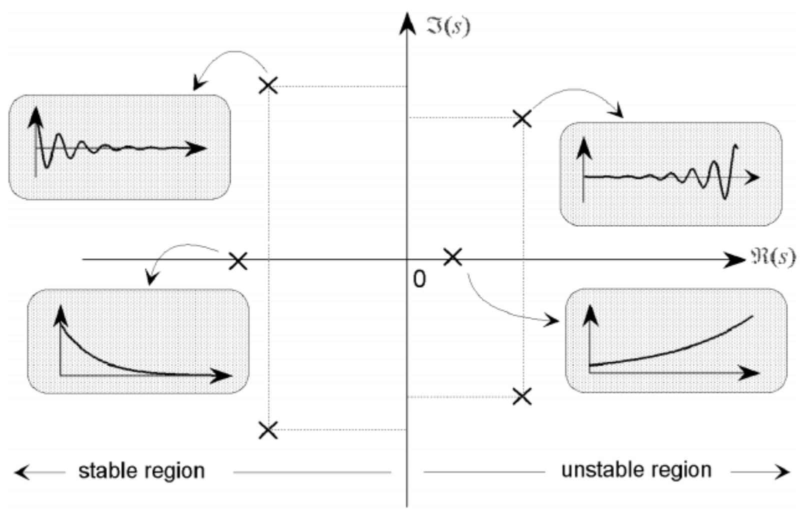
$$D(s) = s^2 + 2s + 5 \rightarrow p_{1/2} = \frac{-2 \pm \sqrt{2^2 - 4 \times 5}}{2} = -1 \pm j2 \tag{4.46}$$

$$G(s) = \frac{N(s)}{D(s)} = 3\frac{(s - (-4))}{(s - (-1 + j2))(s - (-1 - j2))} \tag{4.47}$$

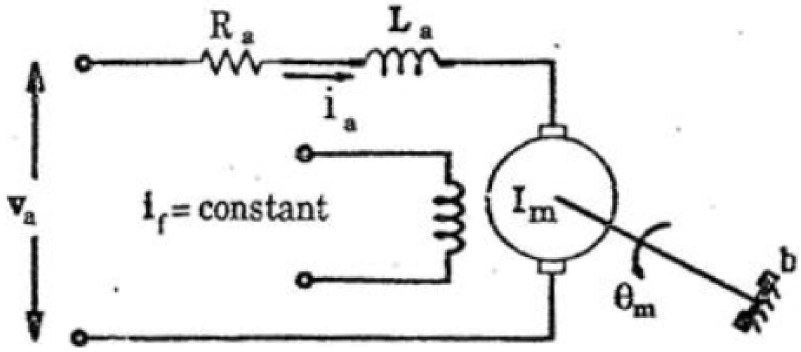


4.5.2 s-Plane

We can see how the time domain representation changes as we move around the s-plane. The imaginary axis corresponds to the sinusoidal component, and the real axis corresponds to the exponential.



4.6 Open Loop Motor Speed Control



For a DC motor connected to an inertial load, there is a combination of electrical and mechanical system equations determining the transfer function for angular position with respect to voltage.

$$\frac{\Theta_m(s)}{v_a(s)} \tag{4.48}$$

First consider the voltage balance:

$$v_a = R_a i_a + L_a \frac{\mathrm{d}i_a}{\mathrm{d}t} + e_m \tag{4.49}$$

Where e_m is the back e.m.f of the motor. We can also write the above in Laplace:

$$V_a(s) = R_a I_a(s) + L_a s I_a(s) + E_m(s) \tag{4.50}$$

Next consider the torque balance in motor:

$$T(t) = I_m \frac{\mathrm{d}^2\theta_m}{\mathrm{d}t^2} + b\frac{\mathrm{d}\theta_m}{\mathrm{d}t} \tag{4.51}$$

Which gives the following in Laplace:

$$T(s) = I_m s^2 \Theta_m + bs\Theta_m(s) \tag{4.52}$$

The electrical and mechanical sides are connected by motor constant which describe the characteristics of the motor (and given by manufacturer).

$$e_m = K_e \frac{\mathrm{d}\theta_m}{\mathrm{d}t} \tag{4.53}$$

$$T = K_t i_a \tag{4.54}$$

Or in the Laplace domain:

$$E_m(s) = K_e s\Theta_m(s) \tag{4.55}$$

$$T(s) = K_t I_a(s) \tag{4.56}$$

Combining this with our previous equations:

$$V_a(s) = R_a I_a(s) + L_a s I_a(s) + E_m(s) \tag{4.57}$$

$$T(s) = I_m s^2 \Theta_m(s) + bs\Theta_m(s) \tag{4.58}$$

We arrive at the equation of voltage with respect to angular position:

$$V_a(s) = \frac{R_a b}{K_t}\left[s(\tau_a s+1)(\tau_m s+1)\right]\Theta_m(s) + K_e s\Theta_m(s) \tag{4.59}$$

"Armature" time constant related to electrical side:

$$\tau_a = \frac{L_a}{R_a} \tag{4.60}$$

Motor time constant related to mechanical side:

$$\tau_m = \frac{I_m}{b} \tag{4.61}$$

The transfer function the becomes:

$$\frac{\Theta_m(s)}{V_a(s)} = \frac{\frac{K_t}{R_a b}}{s\left[\tau_m\tau_a s^2 + (\tau_m+\tau_a)s + \left(\frac{K_e K_t}{R_a b}+1\right)\right]} \tag{4.62}$$

Which looks very complicated, but we can make some simplifying assumptions, which quickly make things simple again. The electrical time constant is normally very small when compared to the mechanical one, as manufacturers try to keep the resistance and inductance low, so we can neglect τ_a. The result, then simplifies to the following equation:

$$\frac{\Theta_m(s)}{V_a(s)} = \frac{K}{s(\tau s+1)} \tag{4.63}$$

$$K = \frac{K_a K_t}{(bR_a + K_t K_b)} \tag{4.64}$$

Where K is the remaining leftover constants. This looks **much** simpler now, but we can take a step further, rather than consider the output position, lets look at the transfer function with respect to speed ω

$$\frac{\omega_m(s)}{V_a(s)} = \frac{\mathrm{d}\left(\frac{\Theta_m(s)}{V_a(s)}\right)}{\mathrm{d}t} = s\left(\frac{K}{s(\tau s+1)}\right) \tag{4.65}$$

$$\frac{\omega_m(s)}{V_a(s)} = \frac{K}{(\tau s+1)} \tag{4.66}$$

This is now just a first order system.

Part II

Instrumentation

Chapter 5

General Measurement Systems

A measurement system has to be devised such that the relationship between the real value of a variable and the value actually measured is unambiguously known. For example, when placing a weight on a scale, we must know the relationship between the **true** weight and the **measured** weight. The measurement system must allow:

- Easy interpretation of the measured data

- Provide high degree of confidence

5.1 Transducers

A transducer converts the sensed variable into a detectable signal form. Sometimes the device changes a mechanical quantity into a change in an electrical quantity. For example, a strain gauge converts a change in strain in the specimen to a change in electrical resistance in the gauge. Another example is a thermometer which converts thermal expansion of the liquid (due to a rise in temperature) into a mechanical translation.

5.2 Signal condition circuits

This takes the transducer signal and can convert, compensate or manipulate it into a more usable electrical quantity. This may include filters, compensators, modulators, demodulators, integrators or differentiators. For example, a Wheatstone bridge used with the strain gauge converts the change in the electrical resistance of the gauge to a change in voltage.

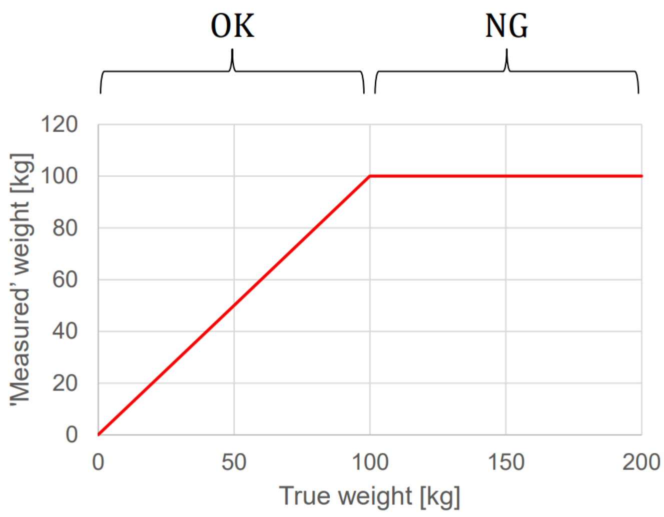


Figure 5.1: The scales work well up to 100kg, however it may not be able to measure heavier weights effectively.

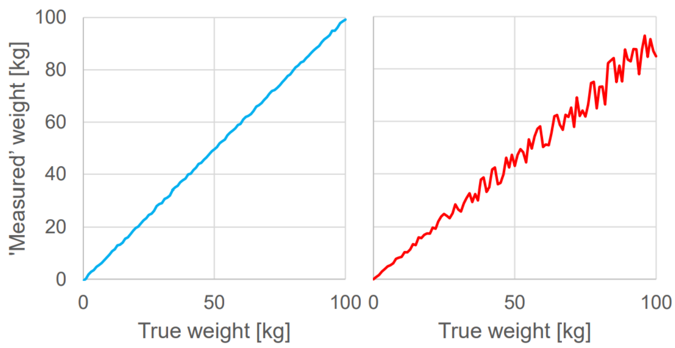


Figure 5.2: On the red graph we can see the effect that noise has on our measurement - a source of an inaccuracy. This must be dealt with to have an effective measurement system.

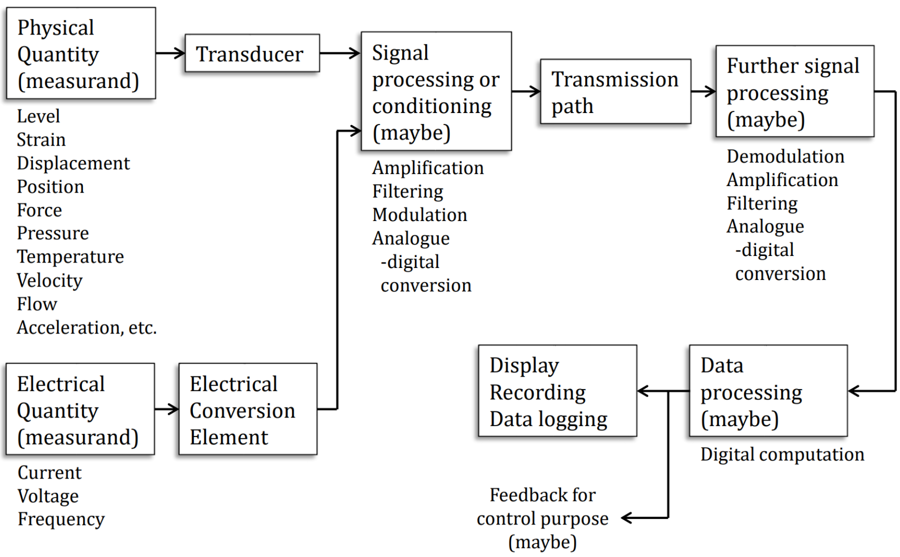


Figure 5.3: Block diagram to show a general measurement system.

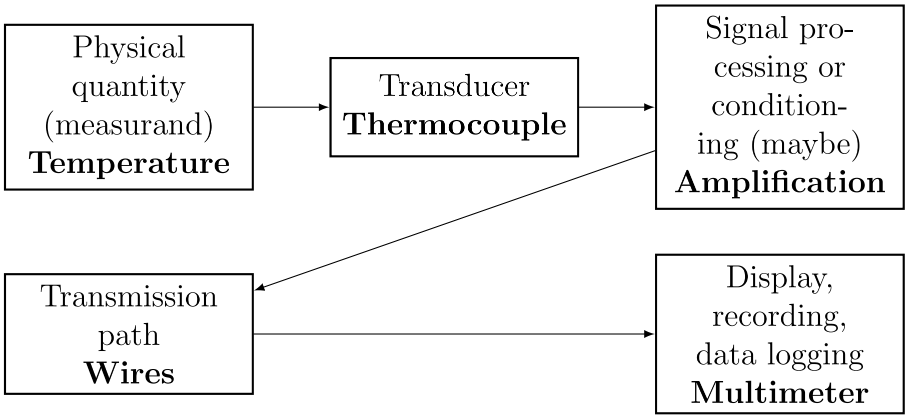


Figure 5.4: How the workflow of a measurement can be constructed.

5.3 Amplifiers

They are required in systems when the output from the transducer-signal conditions is small. Gains of 10 - 1000 are used to increase the levels of the signal, typically a millivolt or less, to what is compatible with the voltage-measuring devices used in the system. A negative-feedback amplifier circuit is an example of an amplifier.

5.4 Output stage

Generally, it is a voltage measurement device that is used to display the measurement in a form that can be read and interpreted. For example, digital voltmeters, self-balancing potentiometers, oscilloscopes, chart recorders and magnetic tape recorders.

5.5 Feedback-control stage

Used when the measurement system is employed in process control. The signal from the measurement system is compared with the command signal that reflects the required value of the quantity in the process. The process-controller forms the difference between these two and produces and error signal. The error signal is then used to automatically adjust the process. For example, the float system in a toilet (called a ballock) to control the water supply in the tank.

5.6 Calibration

How do we know whether one metre is one metre? We compare it with a reference, which is supposed to be reliable. This act of checking is called calibration. Our references are the measuring **standards**.

5.7 Standards

There are seven fundamental quantities of the International Measuring System.

- Length m - metre
- Time s - second
- Mass kg - kilogram
- Temperature K - kelvin
- Electric Current A - ampere
- Luminous Intensity - cd - candela
- Amount of Substance - mol - mole

Units and standard for all other quantities are derived from these.

We can look at how the definition of the metre has changed over the past 2 centuries for some insight into the accuracy of our reference.

- 1792 - a ten millionth of the quarter of a meridian ($0.5 - 0.1$mm of uncertainty)
- 1889 - a platinum-iridium bar at melting point of ice ($0.2 - 0.1$µm of uncertainty)
- 1983 - the length of the path travelled by light in vacuum during a time interval of 1/299792458 of a second (0.1nm of uncertainty)

Here are how the other fundamental quantities are defined.

- Metre - Length of the path travelled by light in a vacuum in 1/299792458 of a second
- Time - Duration of 9192631770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom
- Mass - ~~Equal to the mass of the International Prototype of the Kilogram~~
  - Equal to h / (Metre standard^2 / Time standard) ... since 2019 h: Plank constant (in $e = hf,\ h = 6.626 \times 10^{-34}\mathrm{kg\,m^2\,s^{-1}}$)

- Temperature - 1/273.16 of the thermodynamic temperature of the triple point of water (exactly 0.01 °C)
- Electric current - The flow of electric charges through a surface at the rate of one coulomb per second

The fundamental standards are protected by such organisations such organisations as the National Institute of Standard and Technology (NIST) in the US. They set up the National Reference Standards and below these in accuracy, the Working Standards (and more below these).

5.8 Units and dimensions

- Dimension - a physical variable that is used to describe some aspect of a physical system
- Unit - a measure of a dimension. Primary standards are used to form the exact definition of a unit

SI units are now commonly used worldwide. The fundamental dimensions are:

- Length [L]
- Mass [M]
- Force [F] ([MLS^{-2}])
- Time [τ]
- Temperature [T]

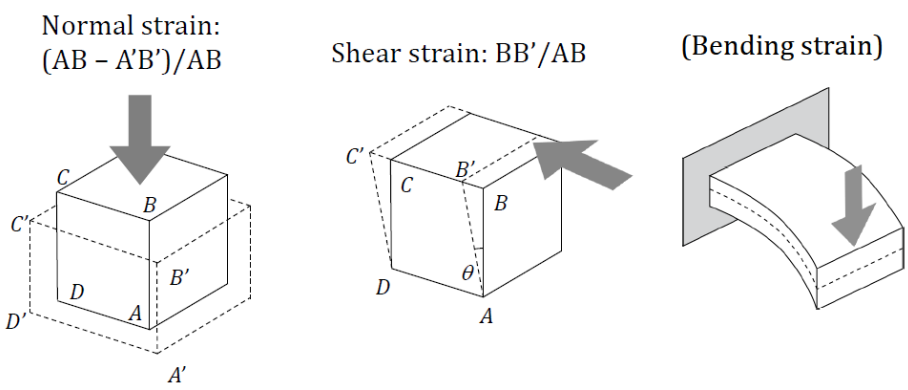
Chapter 6

Strain Gauges

6.1 Strain gauges

6.1.1 Strain

The changes in the value of a dimension of a body divided by the original value of the dimension is the relative change in the dimensions.



Information about strain is required in many engineering situations: an aircraft in flight, support pillars and spans of bridges, etc.

6.1.2 Stress

Density of the reactive (internal) forces distributed throughout the body (force per unit area), in response to external force. Types of stress:

- Tensile/compressive stress
- Shear stress

- Bending stress

Elastic modulus:

$$E = \frac{\text{stress}}{\text{strain}} \tag{6.1}$$

- Young's modulus - normal strain (and if relationship is linear)
- Shear modulus - shear strain
- Bulk modulus - volumetric strain

6.1.3 Strain gauges

A device which experiences a change of electric resistance when it is strained. A strain gauge is combined with other electrical components to obtain an electric voltage or current representing tensile, compressive or bending strain, together wit the means of displaying or recording its value. The total resistance of a block of conducting material of uniform cross section A and length l:

$$R = \rho \frac{l}{A} \tag{6.2}$$

Where ρ is resistivity with units $\Omega\,\mathrm{m}$.

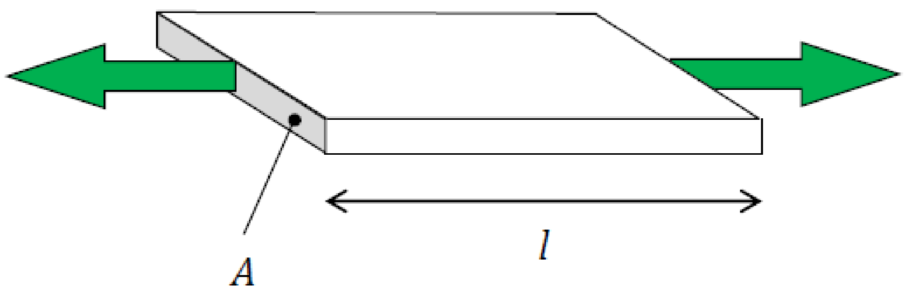


Figure 6.1: When the block is subjected to a tensile stress, its length will increase and cross-sectional area will decrease, both of which cause the resistance to increase according to equation (6.2)

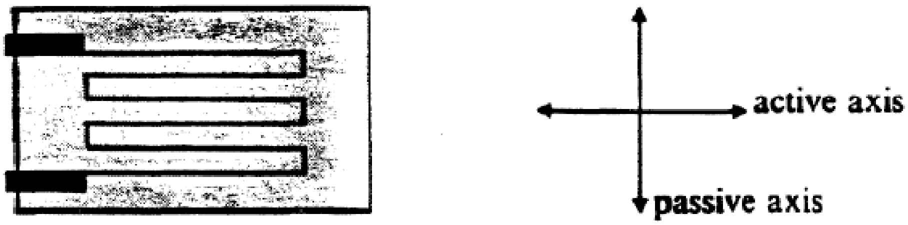
The resistivity of the material will also change because of the **piezo-resistive effect** (increase with tension and decrease with compression). The resistance of the block can thus be written as:

$$R' = (\rho + \delta\rho)\frac{l + \delta l}{A - \delta A} \tag{6.3}$$

Note: compressive stress will result in a decrease in total resistance.

6.1.4 Simple wire strain gauge

A long fine wire is folded to fit in a small area and then mounted on a flexible backing sheet, usually paper. When **firmly stuck** to the surface of a much more rigid body, any deformation of this body will cause an identical fractional change of the strain gauge wire. The change of resistance for any strain along the active axis will be much greater than for an equal strain occurring along the passive axis.



Single-active-axis strain gauge

6.1.5 Gauge factor

Defined as the fractional change of resistance of the gauge, divided by the fractional change in the length of the gauge along the active axis:

$$G = \frac{\frac{\Delta R}{R}}{\frac{\Delta l}{l}} \tag{6.4}$$

Since $\frac{\Delta l}{l}$ is the strain e in the body to which the gauge is fixed, this can be rewritten as:

$$\frac{\Delta R}{R} = eG \tag{6.5}$$

Most gauges have a G between **1.8 to 2.2**, depending on the gauge material and the magnitude of the piezo-resistive effect.

6.1.6 Foil strain gauge

Most modern strain gauges are formed by rolling out a thin foil of the resistive material and then cutting away parts of the foil by a photo-etching process, to create the required grid pattern.

- Usually supplied with think backing paper as electrical insulation
- Adhesive layer for the fixation to the specimen should be creep-free and allow heat dissipation

Advantages over simple wire strain gauges:

- Larger surface area results in a larger area of adhesion
- Accurate reproducibility due to photo-etching technique
- Small dimensions mean they are good for localised strain, fit well to curvature

6.1.7 How can we measure strain

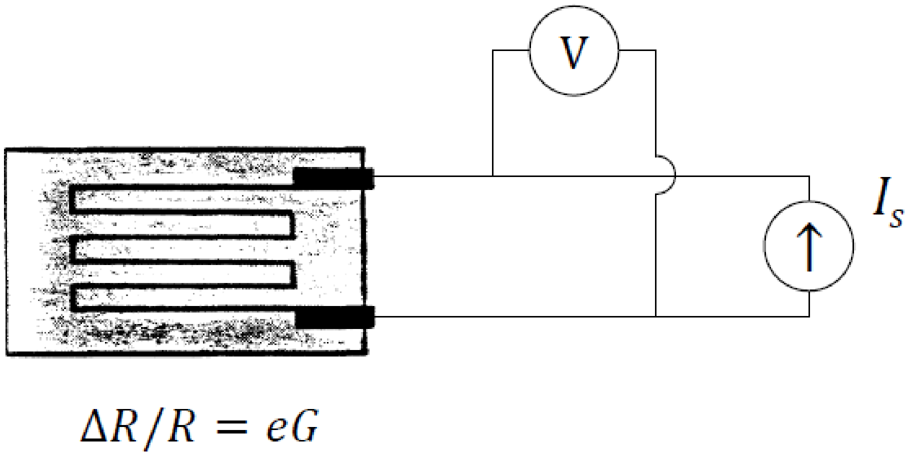


Figure 6.2: Is this good enough?

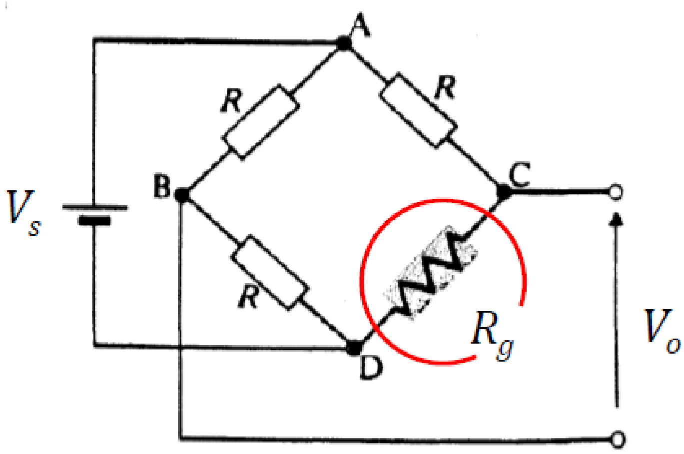
Typically $e = 10^{-3},\ G = 2.1,\ R = 120\Omega,\ I_S < 50\text{mA}$.

$$V = 6\text{V},\ \Delta R = 0.252\Omega \tag{6.6}$$
$$\Delta V = 0.013\text{V} \tag{6.7}$$

The fractional change of the voltage is too small to detect comparing to the 'baseline' voltage ($= RI_S$).

6.2 Wheatstone bridge

Used to convert the change of resistance in the strain gauge into an electrical signal which could be used to indicate the value of strain. With this, **zero strain $\leftrightarrow$ zero output voltage**.



Here,

$$V_C - V_D = \frac{V_S R_G}{R + R_G} \tag{6.8}$$

and then,

$$V_0 = V_C - V_B = \left(V_D + \frac{V_S R_G}{R + R_G}\right) - \left(V_D + \frac{V_S}{2}\right) = \frac{V_S R_G}{R + R_G} - \frac{V_S}{2} \tag{6.9}$$

R is chosen to have the same value as the unstrained gauge, i.e. $R = R_G$ when zero strain $\rightarrow V_0 = 0$ when zero strain. When the gauge is strained, $R_G = R + r$ and the output voltage is

$$V_0 = V_S \frac{R + r}{2R + r} - \frac{V_S}{2} = V_S \frac{r}{2(2R + r)} \tag{6.10}$$

Now using the gauge factor, $\frac{r}{R} = eG$ and the equation can be rewritten:

$$V_0 = V_S \frac{ReG}{2(2R + ReG)} = V_S \frac{eG}{2(2 + eG)} \approx \frac{1}{4} V_S eG \tag{6.11}$$

Because typically $eG < 0.02 << 2$

Sensitivity:

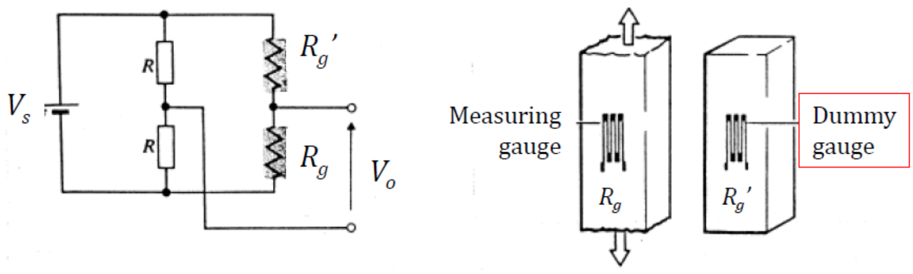
$$\frac{V_0}{e} = \frac{1}{4} V_S G \tag{6.12}$$

6.2.1 Temperature compensation

The output voltage from the strain gauge and the bridge also depends on the temperature. Changes in temperature will cause changes in:

- Dimensions of the specimen and hence the gauge due to thermal expansion
- Dimensions of the gauge itself (particularly thickness)
- Resistivity of the gauge material

Compensated with a dummy gauge:



The output voltage of the bridge is:

$$V_0 = V_S \frac{R_G}{R_G + R'_G} - \frac{V_S}{2} \tag{6.13}$$

Substituting for R_G and R'_G, we have:

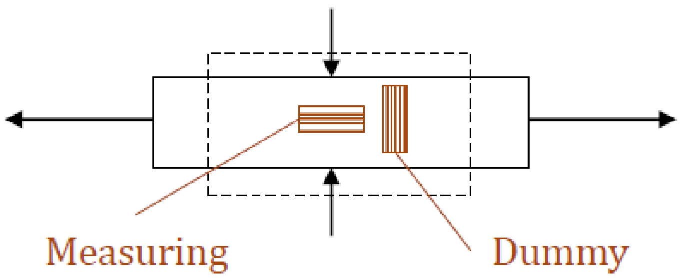
$$V_0 = V_S \frac{R(1+x)1+y}{R(1+x)1+y+R(1+y)} - \frac{V_S}{2} \tag{6.14}$$

$$\frac{V_0}{V_S} = \frac{R(1+x)1+y}{R(1+x)1+y+R(1+y)} - \frac{1}{2} \leftarrow \text{no influence by y} \tag{6.15}$$

However, for this you must find the specimen of the same material and keep them under the same temperature - are these easy enough? The dummy gauge could be placed on the same member as the measuring gauge with its active axis in the direction normal to that of the strain.

6.2.2 Effect of Poisson's ratio on a dummy gauge

Consider both measuring and dummy gauges are attached to the same specimen but perpendicular to each other.



Deformation of the specimen elongates the measuring gauge and shortens the dummy gauge, and the ratio of strains in the two directions is Poisson's ratio, $v = e_{\text{lateral}}/e_{\text{longitudinal}}$, which is between 0.25 to 4 for most materials. Ignoring temperature changes, the measuring gauge and dummy gauge resistances are then

$$R_g = R(1+x) \tag{6.16}$$

$$R'_g = R(1-\nu x) \tag{6.17}$$

The bridge output is:

$$V_0 = V_S \frac{R_g}{R_g + R'_g} - \frac{V_S}{2} \tag{6.18}$$

Substituting for R_g and R'_g, we have:

$$V_0 = V_S \frac{R(1+x)}{R(1+x) + R_g(1-\nu x)} - \frac{V_S}{2} \tag{6.19}$$

$$\frac{V_0}{V_S} = \frac{R(1+x)}{R(1+x) + R_g(1-\nu x)} - \frac{1}{2} \tag{6.20}$$

$$\frac{V_0}{V_S} = \frac{(1+\nu)}{2[2+(1-\nu)x]} \approx \frac{1}{4}(1+\nu)eG \tag{6.21}$$

Because typically $(1-\nu)x << 2$. Comparing to the single gauge, $\frac{V_0}{V_S} = \frac{1}{4}eG$, sensitivity is increased by a factor $(1+\nu)$ due to the effect of Poisson's ratio on the dummy gauge.

6.3 Practical Aspects of Strain Gauge Measurement

6.3.1 Strain gauge rosette

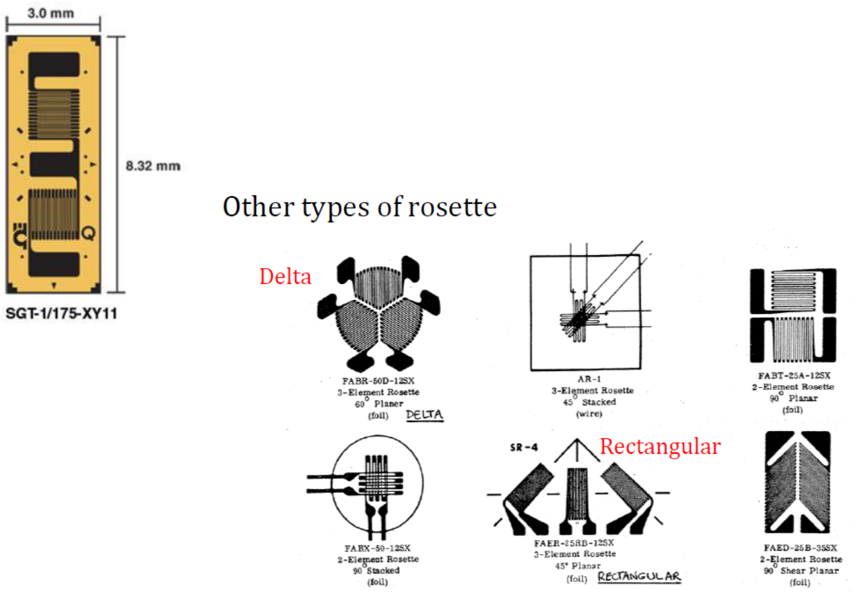
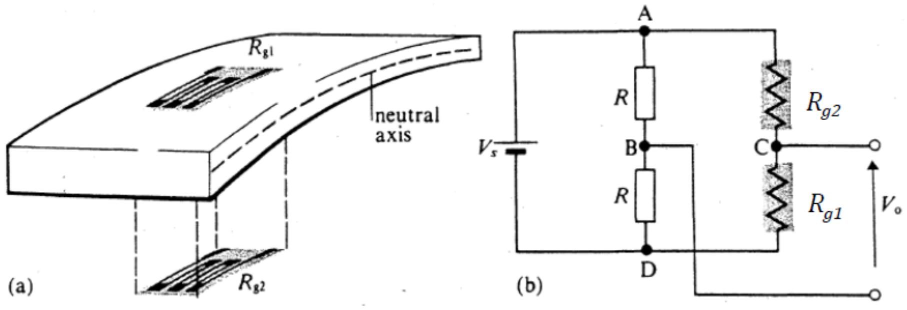


Figure 6.3: An example of a foil strain gauge having two elements with their active axis perpendicular to each other is shown in the top left (Omega Engineering Inc. USA). Some other types of rosette are also shown (figures courtesy of BLH Electronics, Waltham, Mass)

How can be these useful? Strain (stress) in real-world applications is, not uniformly distributed and/or the principal stress axis is unknown. We use strain gauge rosettes to measure strain in three orientations. Understand complex stress/strain state (principal stress calculation is a topic of MECH0013 in term 2).

6.3.2 Bending Strain

To measure the strain of a member in which only bending is occurring, i.e. no additional tensile or compressive stresses present, we attach two strain gauges one on either side of its neutral axis with their actives axes along the length of the member.



If the gauges are equidistant from the neutral axis, the tensile strain imposed on one will be equal to the compressive strain in the other, and t he change of the resistance is the same magnitude but in opposite directions. Similarly to the previous cases, consider the fractional change of resistance x and then output voltage.

$$V_0 = V_S \frac{R_{g1}}{R_{g1} + R_{g2}} - \frac{V_S}{2} \tag{6.22}$$

$$= V_S \frac{R(1+x)}{R(1+x) + R(1-x)} - \frac{V_S}{2} = \frac{1}{2} V_S x \tag{6.23}$$

Since $x = eG,\ V_0 = \frac{1}{2} V_S eG$ is a measure of bending occurring in the body.

6.3.3 Tensile or compressive strain in a bending member

The arrangement shown below allows to measure any tensile or compressive strain in the member, while ignoring any bending strain.

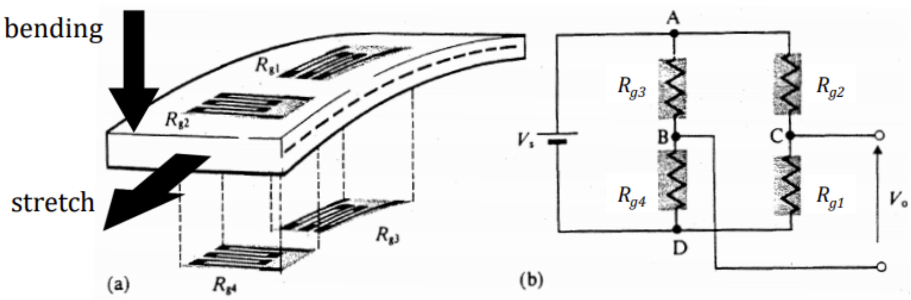


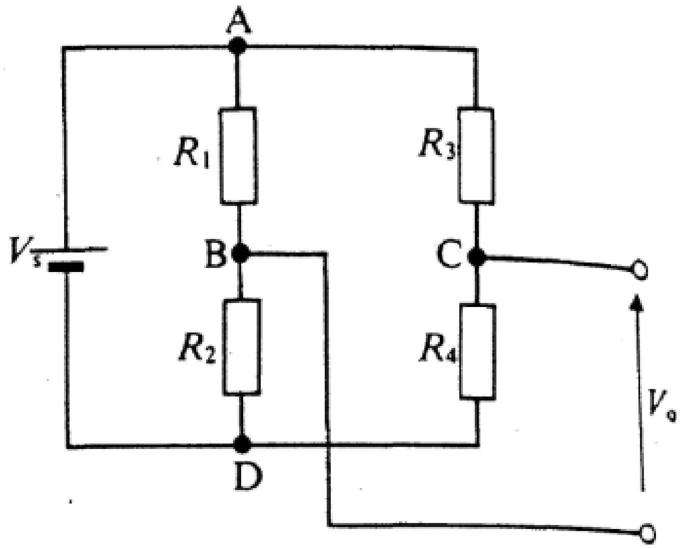
Figure 6.4: Variation of gauge resistance and output voltage in the configuration above.

	Rg_1	Rg_2	Rg_3	Rg_4	V_0
Under pure stretch	++	-	++	-	++
Under pure bending	++	-	- -	+	±0

6.3.4 Bridge balancing

In reality, resistance in the arms of the Wheatstone bridge slightly differ in value because of:

- Manufacturing variations
- Temperature differences between gauges and resistors
- Static strain occurring in a member



If the bridge is perfectly balanced:

$$\frac{R_4}{R_3 + R_4} = \frac{R_2}{R_1 + R_2} \tag{6.24}$$

$$\rightarrow \frac{R_3}{R_4} + 1 = \frac{R_1}{R_2} + 1 \rightarrow \frac{R_3}{R_4} = \frac{R_1}{R_2} \tag{6.25}$$

The imbalance needs to be adjusted before the measurement (adjustment of the output voltage to zero) → bridge balancing. One possible way to achieve this is to connect a pot (potentiometer) in series with R_2. The total resistance $R_x = R_1 + R_p$ should be adjusted such that:

$$\frac{R_x}{R_2} = \frac{R'_g}{R_g} \tag{6.26}$$

6.3.5 Semiconductor strain gauge

In some crystalline materials such as germanium and silicon, the piezo-resistive effect is very large. If a slice of such a crystal is used as a strain gauge, very large gauge factors can be obtained, which could range from 100 to 300.

$$\text{Gauge factor: } G = \frac{1}{e}\frac{\delta R}{R} \tag{6.27}$$

The magnitude of the piezo-resistive effect, which determines the sensitivity of the gauge, depends on the minute quantities of impurity introduced into the base material before formation of the crystal.

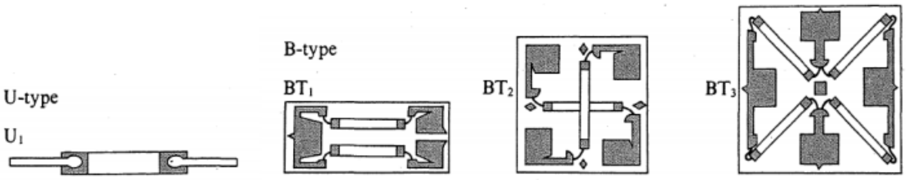


Figure 6.5: Some examples of semiconductor strain gauges

Characteristics of semiconductor strain gauges:

- The material of the strain gauge reaches its elastic limit at about 4000 microstrain (4000×10^{-6}), much smaller than that of metals (typically 20000 microstrain)
- The gauge factor G varies with high strain at high strain levels. i.e. the gauge is not linear
- The gauge factor varies with temperature
- The temperature coefficient of resistivity is large

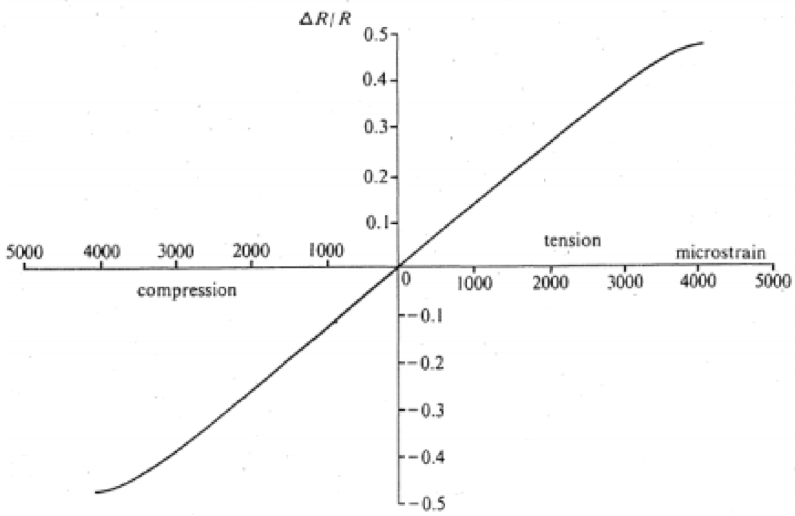
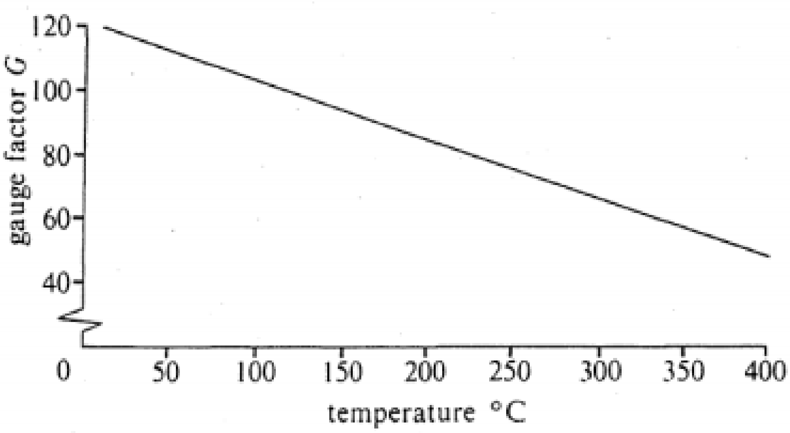
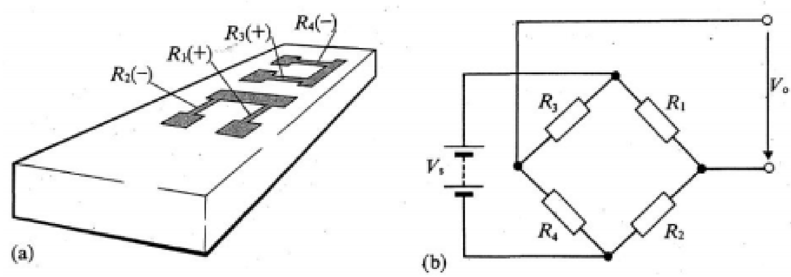


Figure 6.6: Change of resistance ratio with strain for a semiconductor strain gauge



Negative effect of the temperature sensitivity of this type of strain gauge can be significantly reduced by using two gauges each consisting of two crystals connected in series. When two crystals are connected in series (but aligned in parallel), tensile of compressive strain applied to a gauge will produce an increase in the resistance of one crystal and a decrease in the other. Because the base material of each crystal is the same, their temperature coefficients of resistivity are approximately equal. In the configuration below, if the body experiences a tensile strain,

$$R_1 = R(1 + x) \tag{6.28}$$
$$R_2 = R(1 - x) \tag{6.29}$$
$$R_3 = R(1 - \nu x) \tag{6.30}$$
$$R_4 = R(1 + \nu x) \tag{6.31}$$



The bridge output voltage will thus be:

$$V_0 = \frac{1}{2} V_S x(1 + \nu) \tag{6.32}$$

Even if temperature increases, that affects all crystals equally. Also, changes in dimensions will cause $R_1 \uparrow$, $R_2 \downarrow$, $R_3 \uparrow$, $R_4 \downarrow$. Hence, no unbalance occurs.

6.4 Summary

1. Strain gauges utilise variability of resistance due to deformation

2. Strain gauges are used in Wheatstone bridges with various gauge arrangements
3. Bridge balance calculation is an essential part of the module