Introduction to Biomedical Engineering

Section 2: Control theory

Lecture 2.2: 1st and 2nd order systems analysis

POLYTECH
Peter the Great
St.Petersburg Polytechnic
University

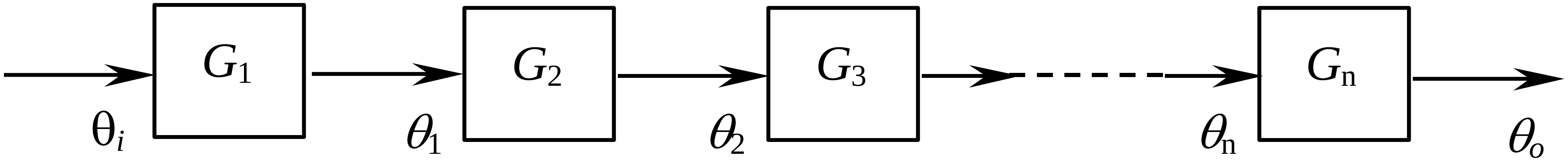
OPEN
POLYTECH

Block Diagram Manipulation

Previously we saw how a feedback loop could be collapsed to a single transfer function. Before moving on to more complex operations, lets recap the basic rules:

What if you had lots of components in series?

Cascading blocks (in series)



(These blocks could be a sensor, into a filter into an amplifier etc…)

These can all be combined into a single block:

$$G_1 = \frac{\theta_1}{\theta_i},\ G_2 = \frac{\theta_2}{\theta_1}\ \text{etc}$$

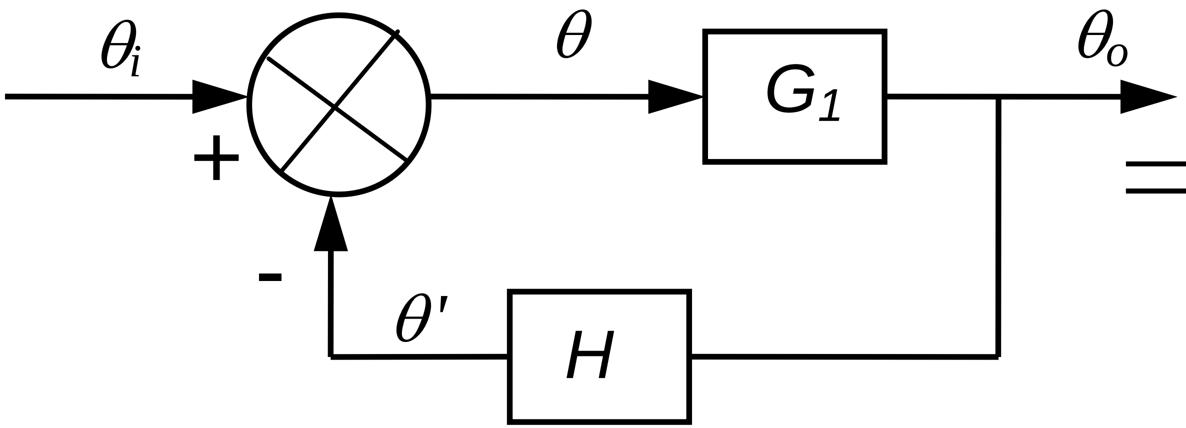
$$\therefore \frac{\theta_o}{\theta_i} = \frac{\theta_o}{\theta_n} \cdot \frac{\theta_n}{\theta_{n-1}} \cdots \frac{\theta_3}{\theta_2} \cdot \frac{\theta_2}{\theta_1} \cdot \frac{\theta_1}{\theta_i}$$

$$= G_n \cdot G_{n-1} \cdots G_3 \cdot G_2 \cdot G_1$$

$$\therefore G = G_1 \cdot G_2 \cdot G_3 \cdots G_n$$

Feedback system

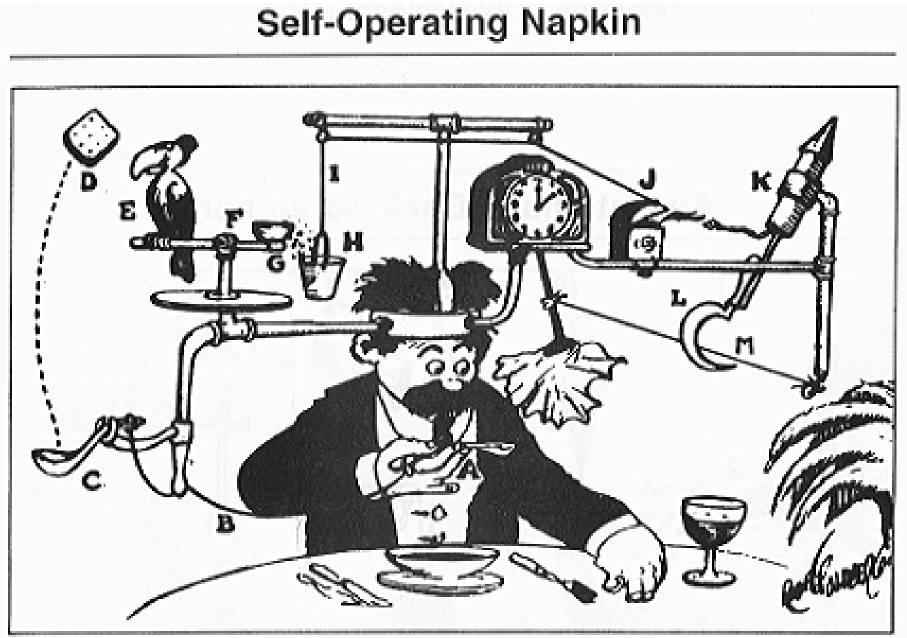
A feedback system with a open loop transfer function G(s) and a feedback gain of H(s) can be combined into a single transfer function F(s)

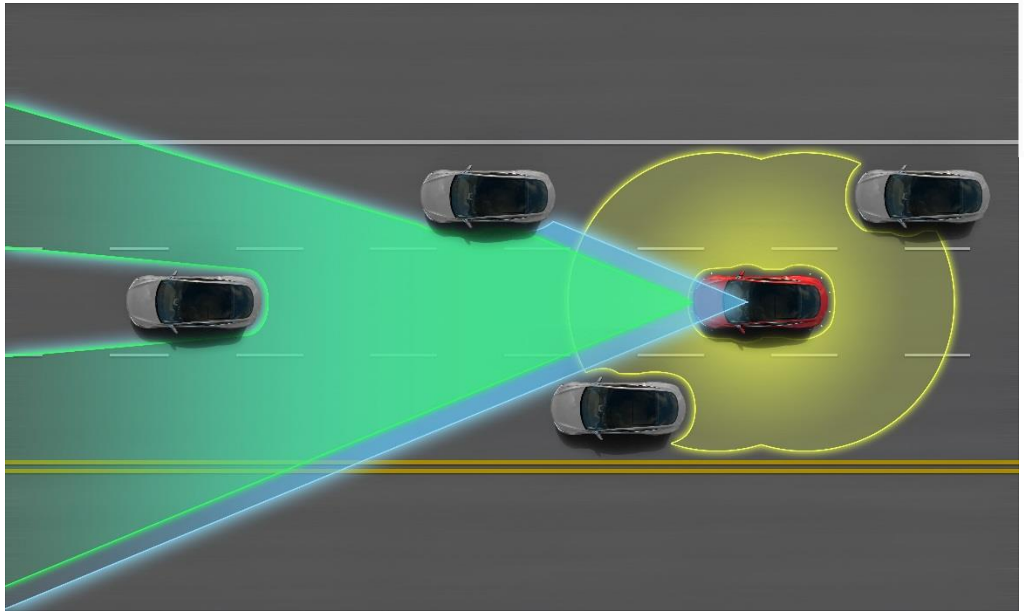


$$\Rightarrow F_1 = \frac{\theta_o}{\theta_i} = \frac{G_1}{1+G_1 H}$$

Nested loops

What if you had loops within loops?

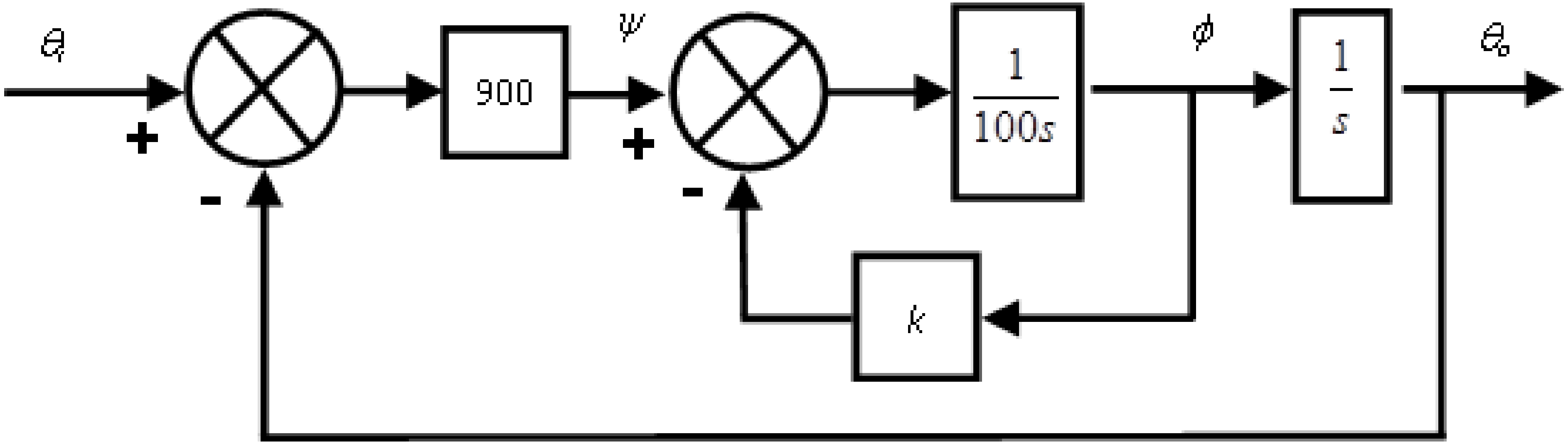




Tesla Autopilot –nested loops for trajectory and speed (*at least!)*

Nested loops

These rules can be applied to more complex systems with nested feedback loops. These occur frequently in control systems, such as a servo system controlling the position of an aileron on a planes wing within the global roll/pitch/yaw control loop.



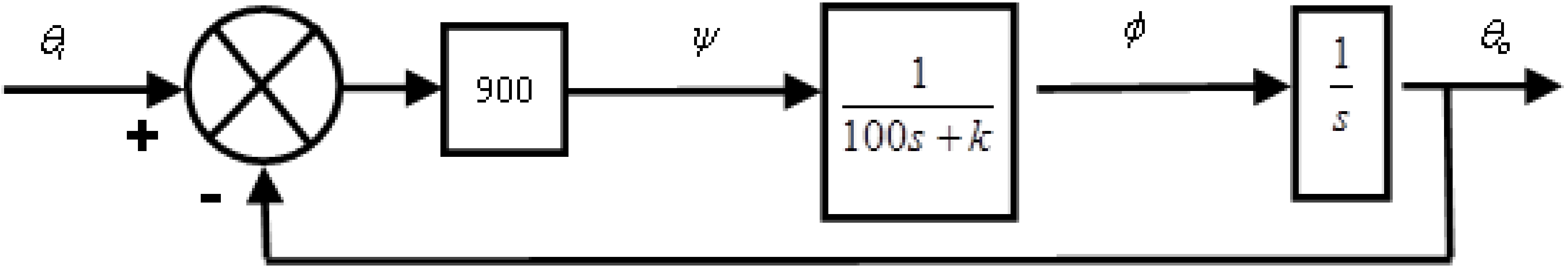
In these cases it is easier to start with the inner loop and work outwards.

If we first call the open loop gain of the inner loop $G = \dfrac{1}{100s}$

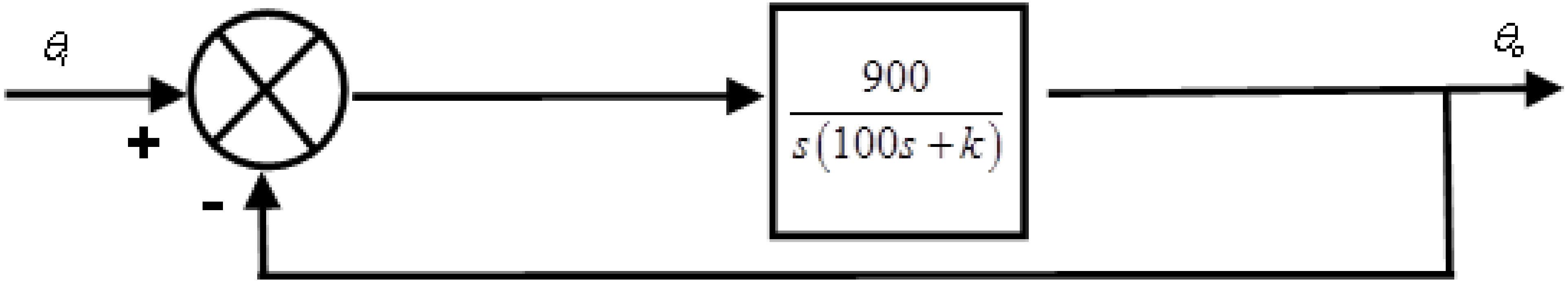
Then, using the feedback system result, the closed loop transfer function becomes:

$$\frac{\varphi}{\psi} = \frac{G}{1+kG} = \frac{1}{100s+k}$$

So the block diagram can be reduced to:



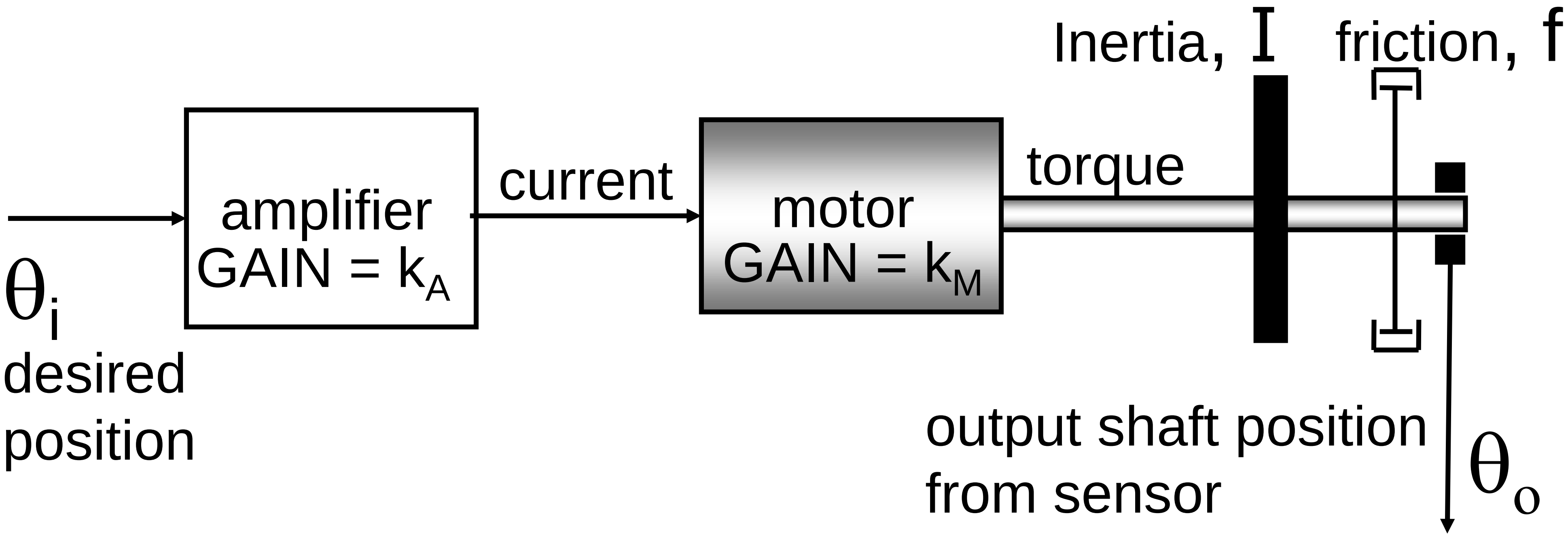
The three blocks in the feed forward path can be combined, so the system is now reduced to a unity feedback system:



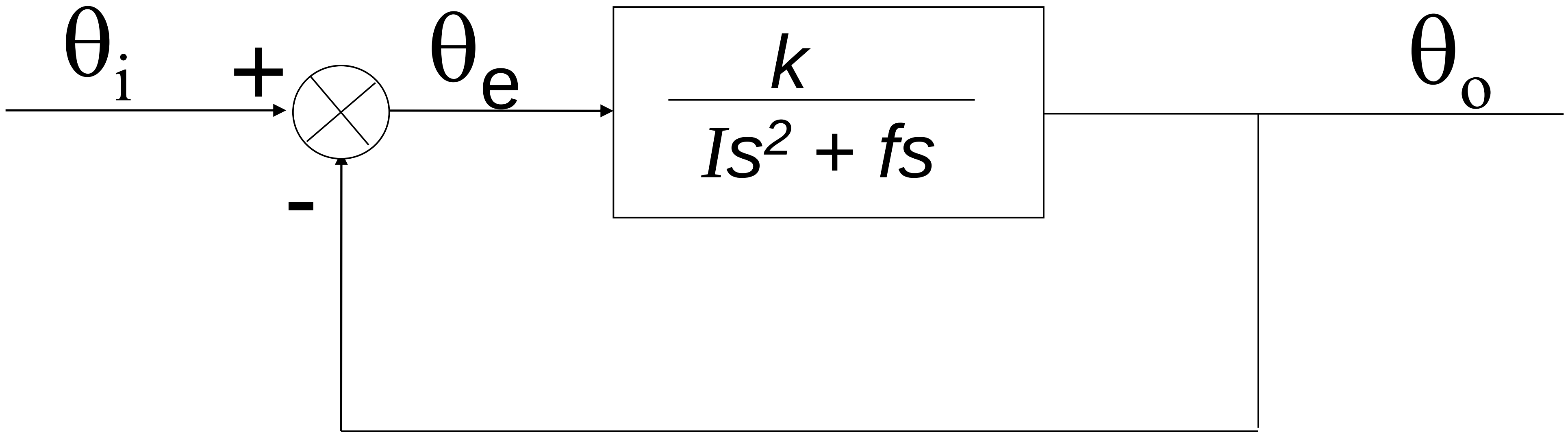
Or as a single transfer function:

$$\frac{\theta_o}{\theta_i} = \frac{9}{s^2 + 0.01ks + 9}$$

Servo as second order system



DC motor servo with position control has a transfer function of a second order system:



$$F(s) = \frac{G(s)}{1+G(s)} = \frac{k/I}{s^2 + f/I\, s + k/I}$$

Where: $\omega_n = \sqrt{\frac{k}{I}}$ $\qquad \zeta = \frac{f}{2\sqrt{kI}}$

Increasing *k* **Increasing *k***

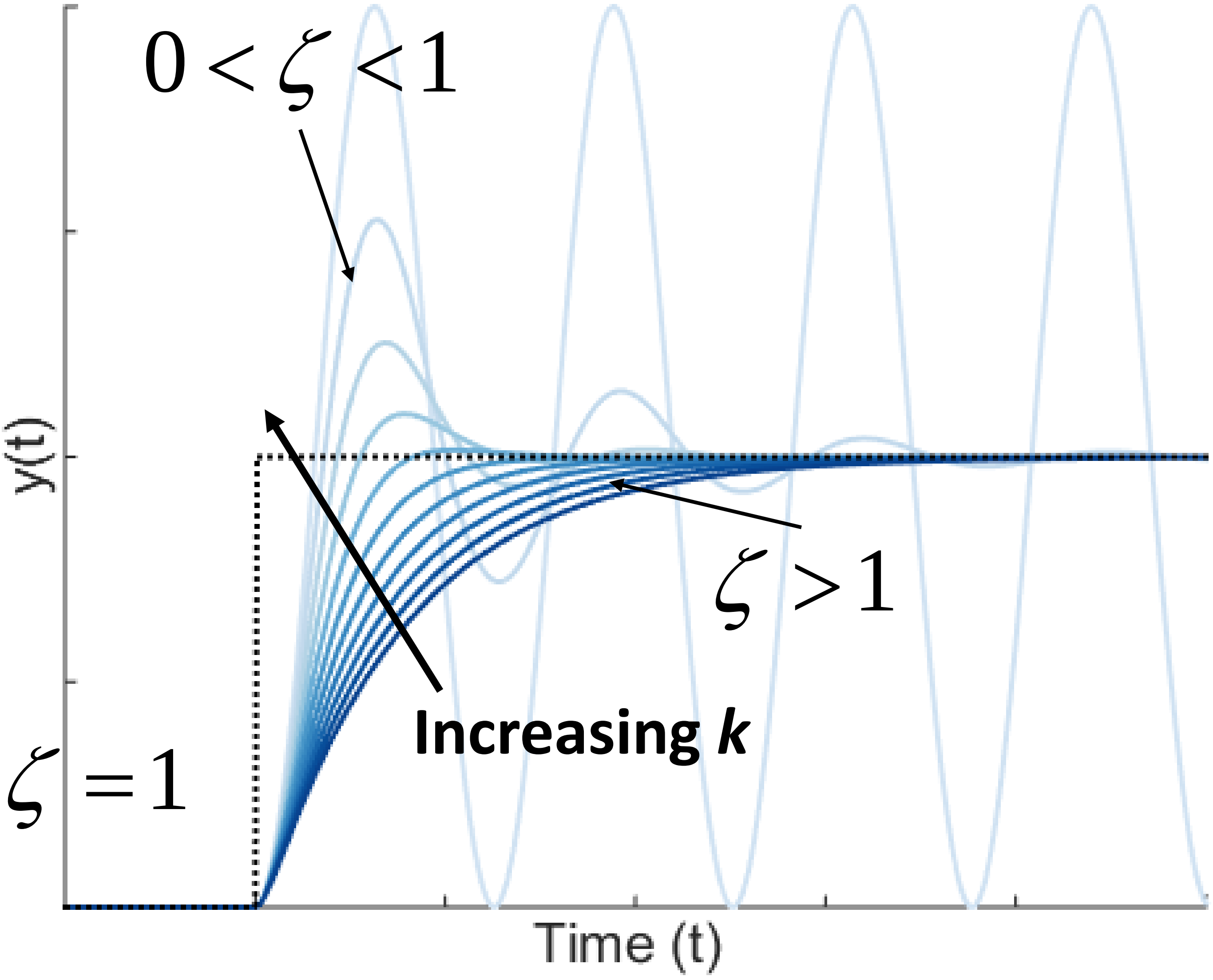
Servo Transfer function

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

Using our knowledge of the characteristics of the transfer function of a second order system we can predict the behaviour of the servo based on the gain *k.*

The roots of the *characteristic equation* determined the response of the system, which was determined by the damping ratio:

$$s = -\zeta\omega_n \pm \omega_n\sqrt{\zeta^2 - 1}$$



$0 < \zeta < 1$
y(t)
$\zeta > 1$
Increasing *k*
$\zeta = 1$
Time (t)

Using the Laplace tables, we are able to obtain the time domain representation for each of these cases

Overdamped	$\zeta > 1$	$s = -\zeta\omega_n \pm \omega_n\sqrt{\zeta^2 - 1}$
Critically Damped	$\zeta = 1$	$s = -\omega_n$
Underdamped	$0 < \zeta < 1$	$s = -\zeta\omega_n \pm j\omega_n\sqrt{1 - \zeta^2}$
Undamped	$\zeta = 0$	$s = \pm j\omega_n$

However, there are other ways of looking at the Transfer Function to obtain a picture of the systems behaviour – often without having to convert back into the time domain.

Transfer functions – Poles and Zeros

As we have seen Transfer Functions allow for important characteristics of the system response to be determined, without having to solve the complete differential equation. Generally, transfer functions can be derived from the differential equations thusly:

Solving this is hard!

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_o y = b_0 x + b_1 \frac{dx}{dt} + b_2 \frac{d^2 x}{dt^2} + b_{m-1} \frac{d^{m-1} x}{dt^{m-1}} + \cdots b_m \frac{d^m x}{dt^m}$$

Where *a* are the output coefficients, and *b* are the input coefficients, combined these *entirely* characterise the system. The Laplace domain transfer function is then given by:

$$H(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0}{a_n s^n + b_{n-1} s^{n-1} + \ldots + a_1 s + a_0}$$

Generally, it is often more convenient to write the transfer function with factorised polynomials in the numerator and denominator.

So first by dividing through by the higher order terms b_m *and* a_m to put into standard form, and then finding the roots of both the numerator and denominator:

$$H(s) = K\frac{(s-z_1)(s-z_2)\dots(s-z_{m-1})(s-z_m)}{(s-p_1)(s-p_2)\dots(s-p_{m-1})(s-p_m)}$$

Where *K* is the gain of the system $K = b_m / a_m$

This term is often neglected as it only *scales* the response, it does not determine its behaviour **Oscillate/decay fast or slow?/Stable?**

Looking at the numerator, the z_i terms are called ***zeros,*** as when $s=z_i$ the numerator and subsequently the transfer function H equals zero

$$H(s) = K\frac{(s-z_1)(s-z_2)\ldots(s-z_{m-1})(s-z_m)}{(s-p_1)(s-p_2)\ldots(s-p_{m-1})(s-p_m)}$$

Looking at the denominator, the p_i terms are called ***poles,*** as when $s=p_i$ the denominator becomes zero, which means the transfer function H goes to infinity (becomes *unbounded*)

Together, K and the poles and zeros *completely define* the transfer function, and thus *completely represent* the system's differential equation

Transfer functions Poles and Zeros Example

As the polynomials in the numerator and denominator are real, the poles and zeros must be either purely real, or in complex conjugate pairs (so the imaginary parts cancel)

Lets take an example linear ODE to demonstrate

$$\frac{d^2y}{dt^2} + 5\frac{dy}{dt} + 6y = 2\frac{dx}{dt} + 1x$$

$$H(s) = \frac{2s+1}{s^2+5s+6} = \frac{1}{2}\frac{s+1/2}{(s+3)(s+2)}$$

So the system has a single real zero at s=-0.5, and two real poles at s=-3 and s=-2

So for the second order/servo case:

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

Where: $a = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}, \mathrm{b} = -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}$

The system has *no* zeros, and a pair of poles, which can be real, imaginary or complex depending upon the damping ratio

Transfer functions in the S Plane

A system is characterised by its poles and zeros as they allow the reconstruction of the input/output differential equation.

It is possible to get a sense of the system dynamics from plotting the poles and zeros graphically on the s –plane.

A pole is commonly represented by a cross (x) and a zero by a circle (○).

Transfer functions in the S Plane

$$s = \sigma + j\omega$$

Recall that, as *s* is a complex number, we can plot the value on a plane, with a real and imaginary axis.

By plotting the locations of the poles and zeros on this plane, we can obtain a considerable amount of information about the response of the system *without* having to take the Inverse Laplace transform.

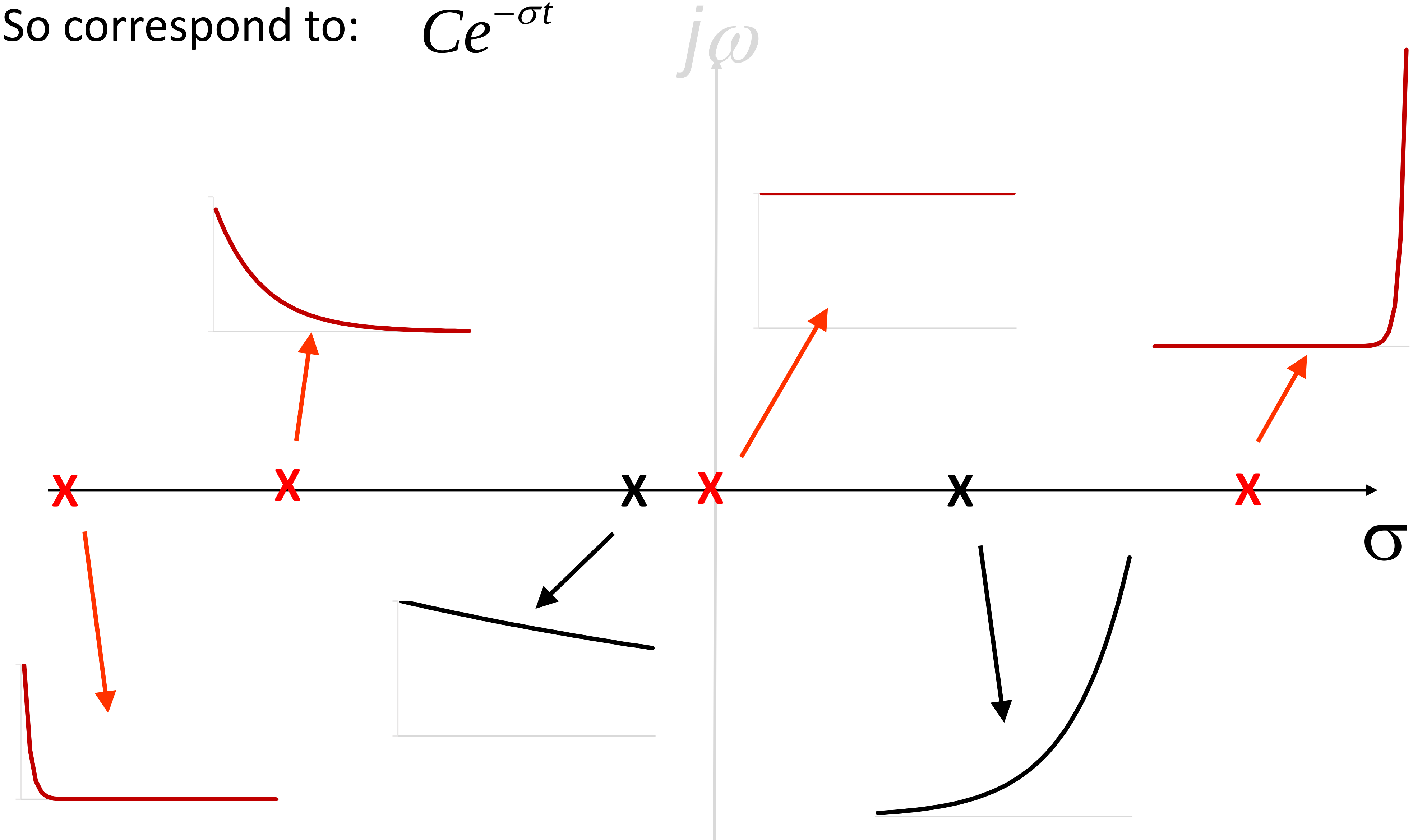
Each pole corresponds to a component of the time domain response, so from the plot it is possible to determine:

- What components exist
- Their relative importance (and possible simplifications)
- How they change with gain

Real Poles – Exponentials

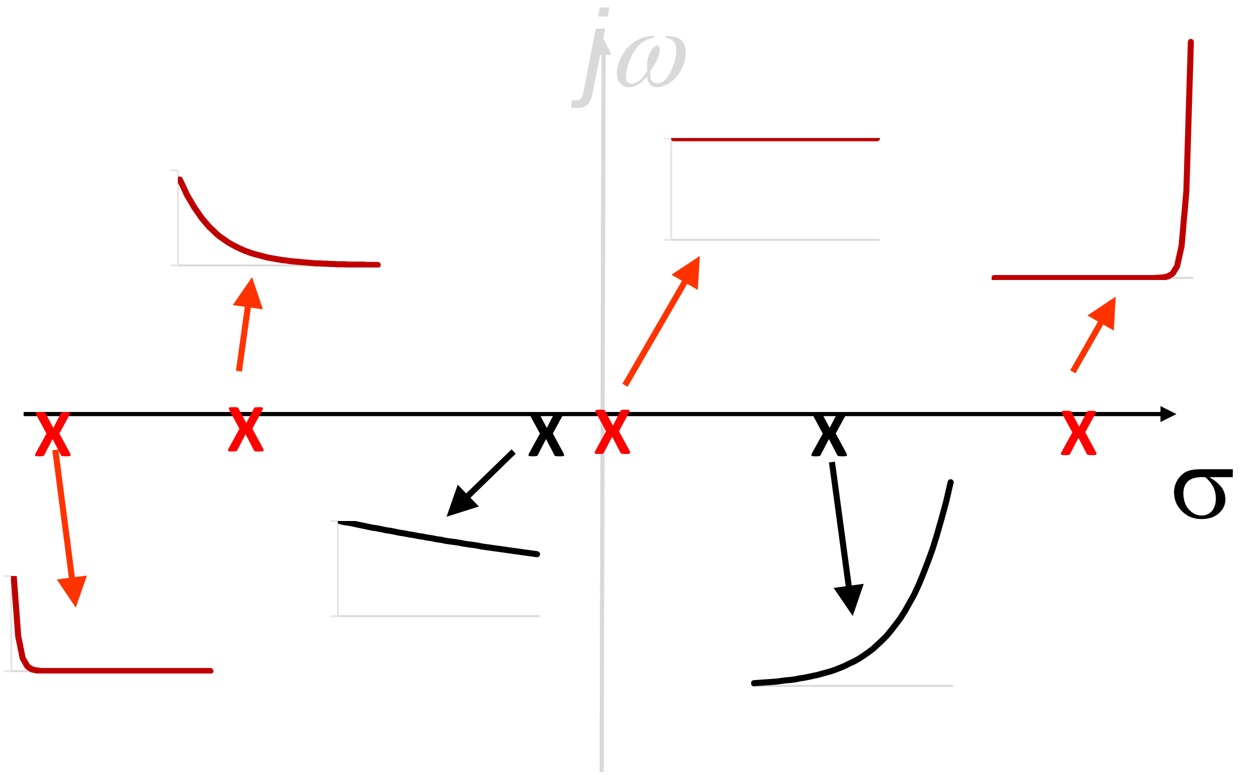
First lets look at the real component only: $s = \sigma$ $\qquad e^{s} \rightarrow e^{-\sigma}$

So correspond to: $Ce^{-\sigma t}$



Real Poles – Exponentials

Poles located on the real axis have an exponential component only, with the pole location determining the rate of decay. Poles close to origin decay slowly, where poles far away decay rapidly.



Poles located in the right hand side exponentially *increase*, whereas poles on the left hand side *decrease.* A pole on the origin is a flat “DC” response

Imaginary Poles – Sinusoids

First lets look at the real component only: $s = \pm j\omega$

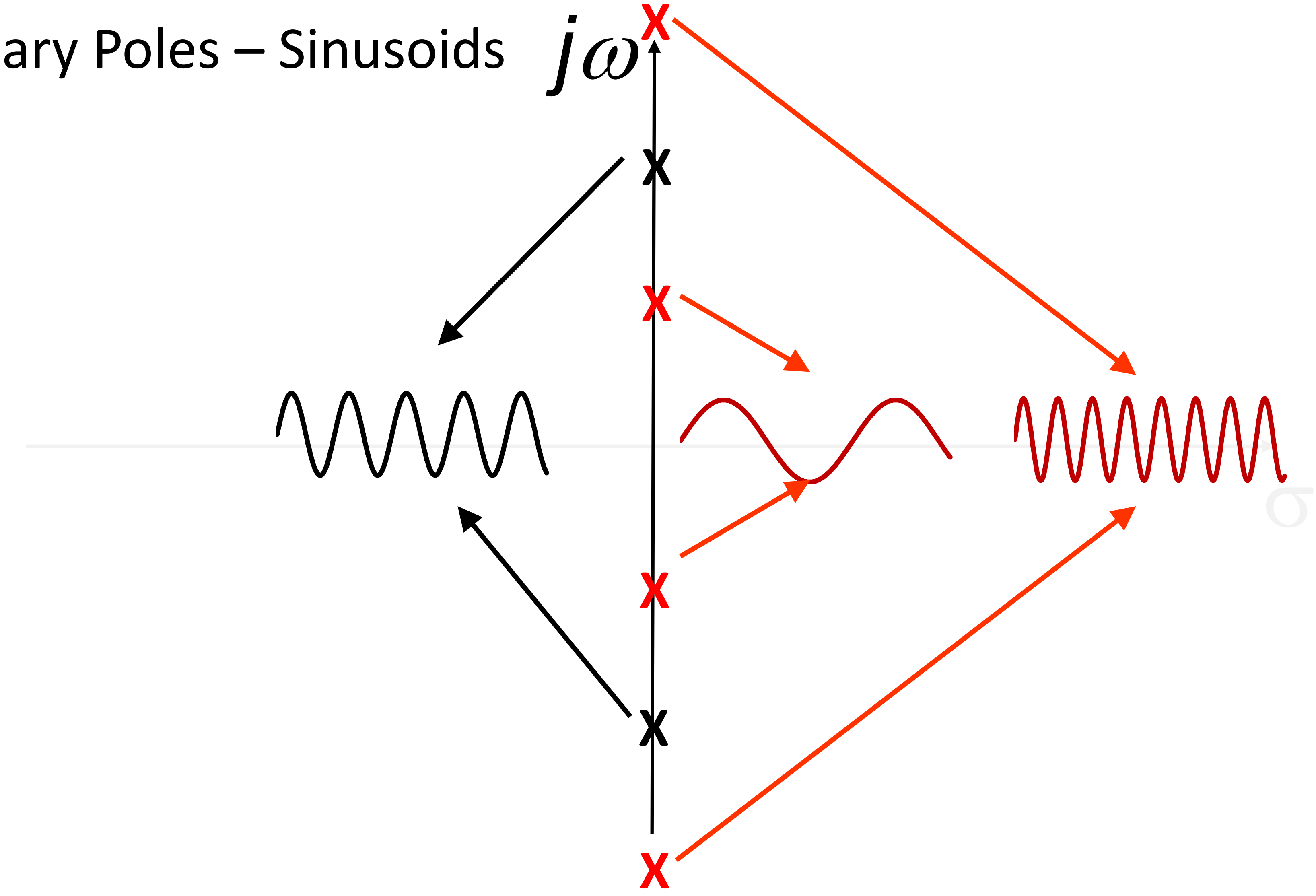
$$e^{s} \rightarrow e^{\pm j\omega}$$

Which, from Eulers formula gives a sinusoidal response

$$\left(e^{jx} = \cos\left(x\right) + j\sin\left(x\right)\right)$$

Poles located on the imaginary axis are in conjugate pairs

Imaginary Poles – Sinusoids



The imaginary poles generate an oscillatory component with a constant amplitude. Poles closer to the origin have a low frequency, which increases the further poles are from the origin

Complex Poles – Combined $s = \sigma \pm j\omega$

A complex conjugate pair of complex poles generate an exponentially decaying sinusoid the form:

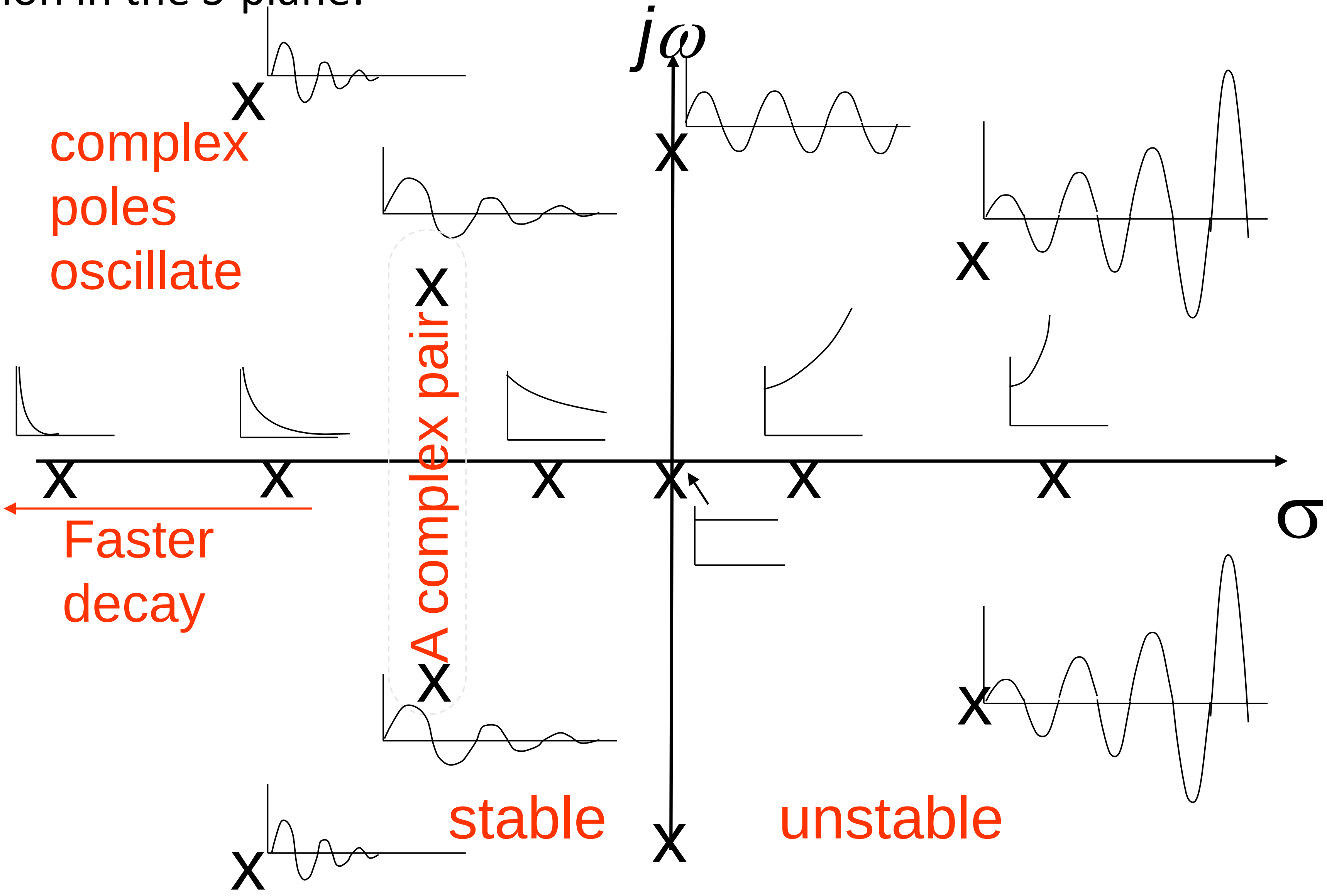
$$Ae^{-\sigma t}\sin\left(\omega t + \phi\right)$$

A and φ are determined by initial conditions, with ω specifying oscillations and σ the rate of decay.

Poles located in the left hand side decay to zero, whereas poles in the right hand increase to infinity, thus making the system ***unstable***

Pole Location and Impulse response summary

The impulse response of each pole in the transfer function depends upon its location in the S-plane:



Second Order System – S plane

As we have seen, the roots of the transfer function of a second order system depends upon the damping coefficient

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

We can plot the poles of the characteristic equation as with any other transfer function

$$s = -\zeta\omega_n \pm \omega_n\sqrt{\zeta^2 - 1}$$

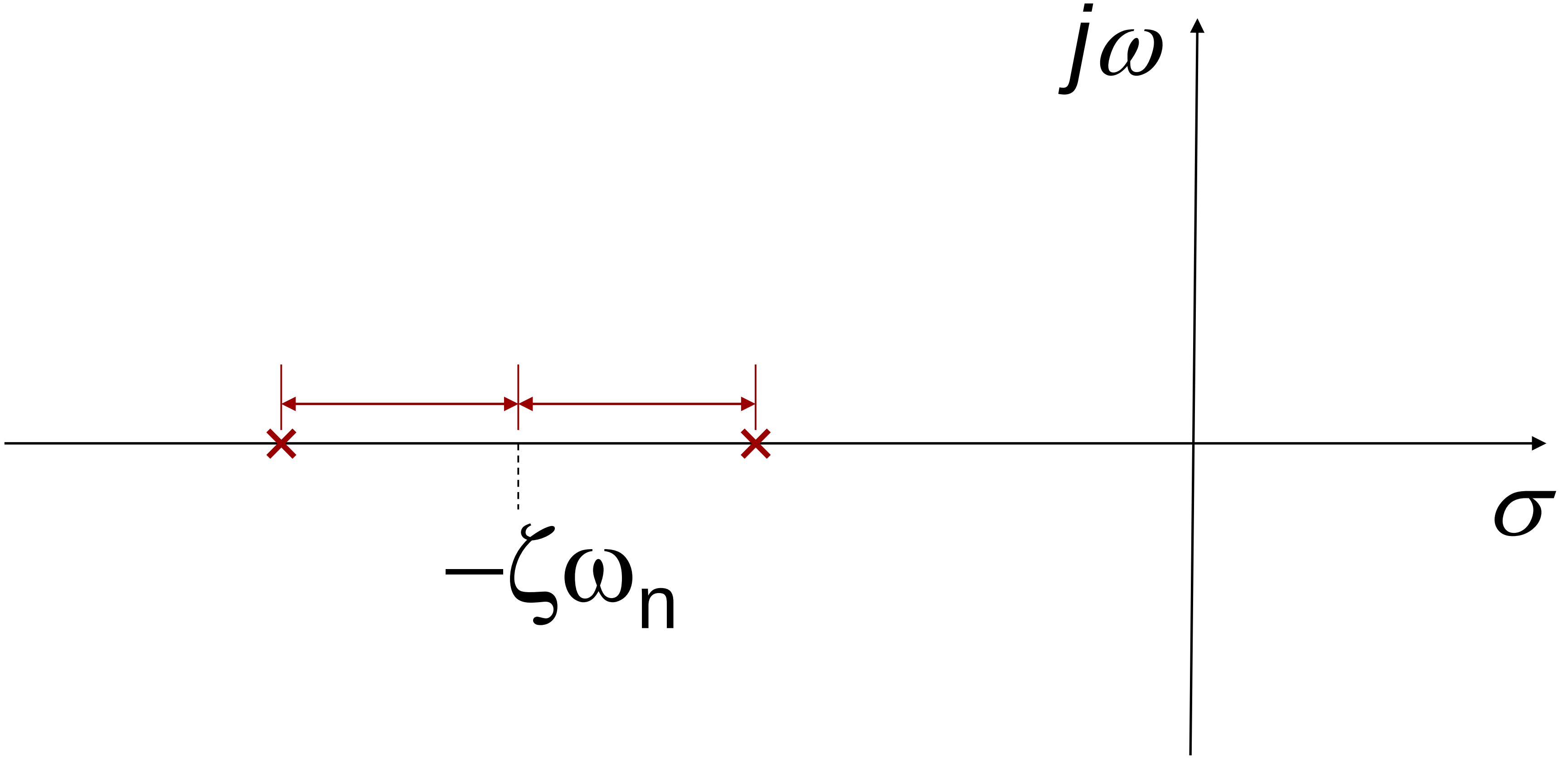
Overdamped $\zeta > 1$

Both roots are real, and thus are on real axis

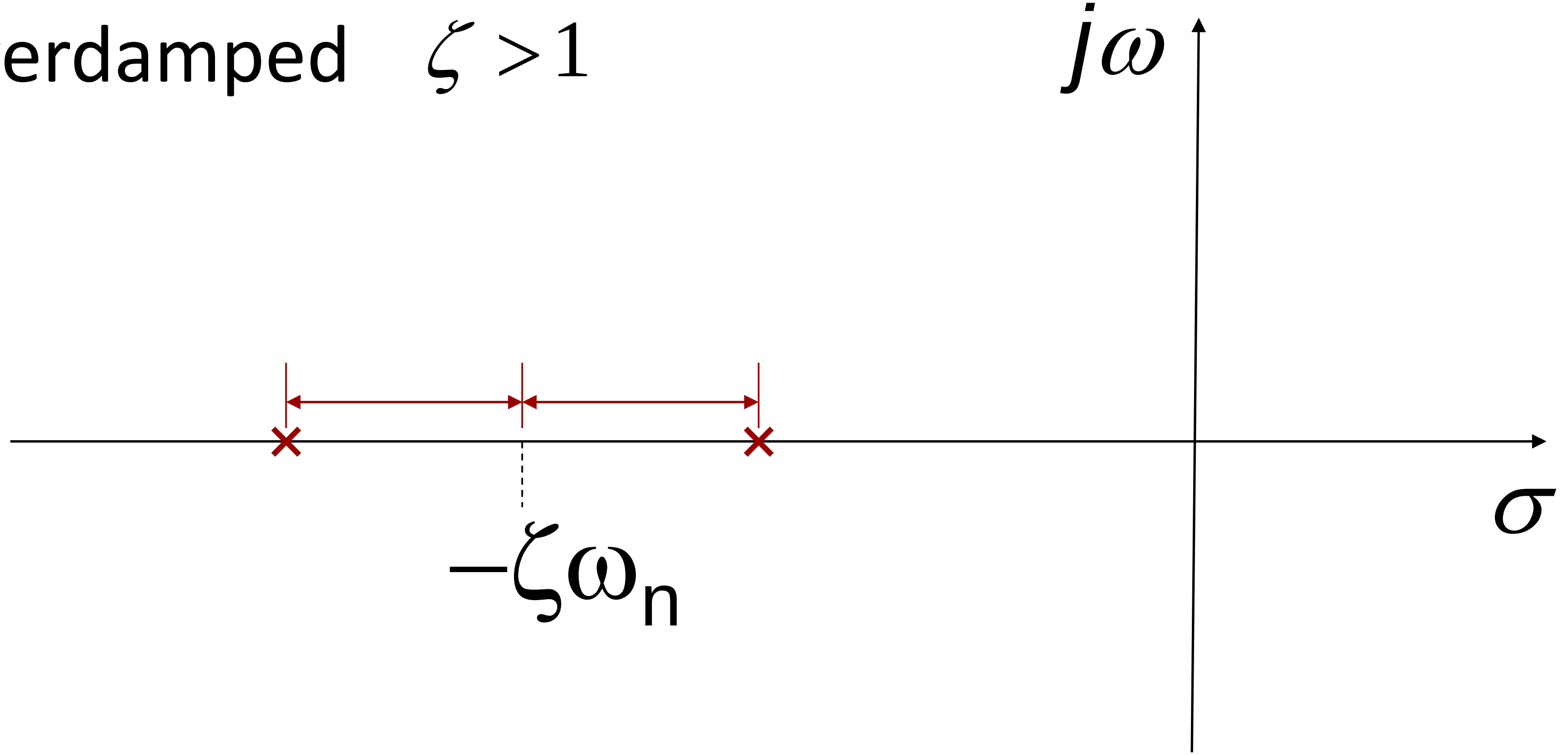
No oscillations in step response

$$s_1 = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}$$

$$s_2 = -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}$$



Overdamped $\zeta > 1$

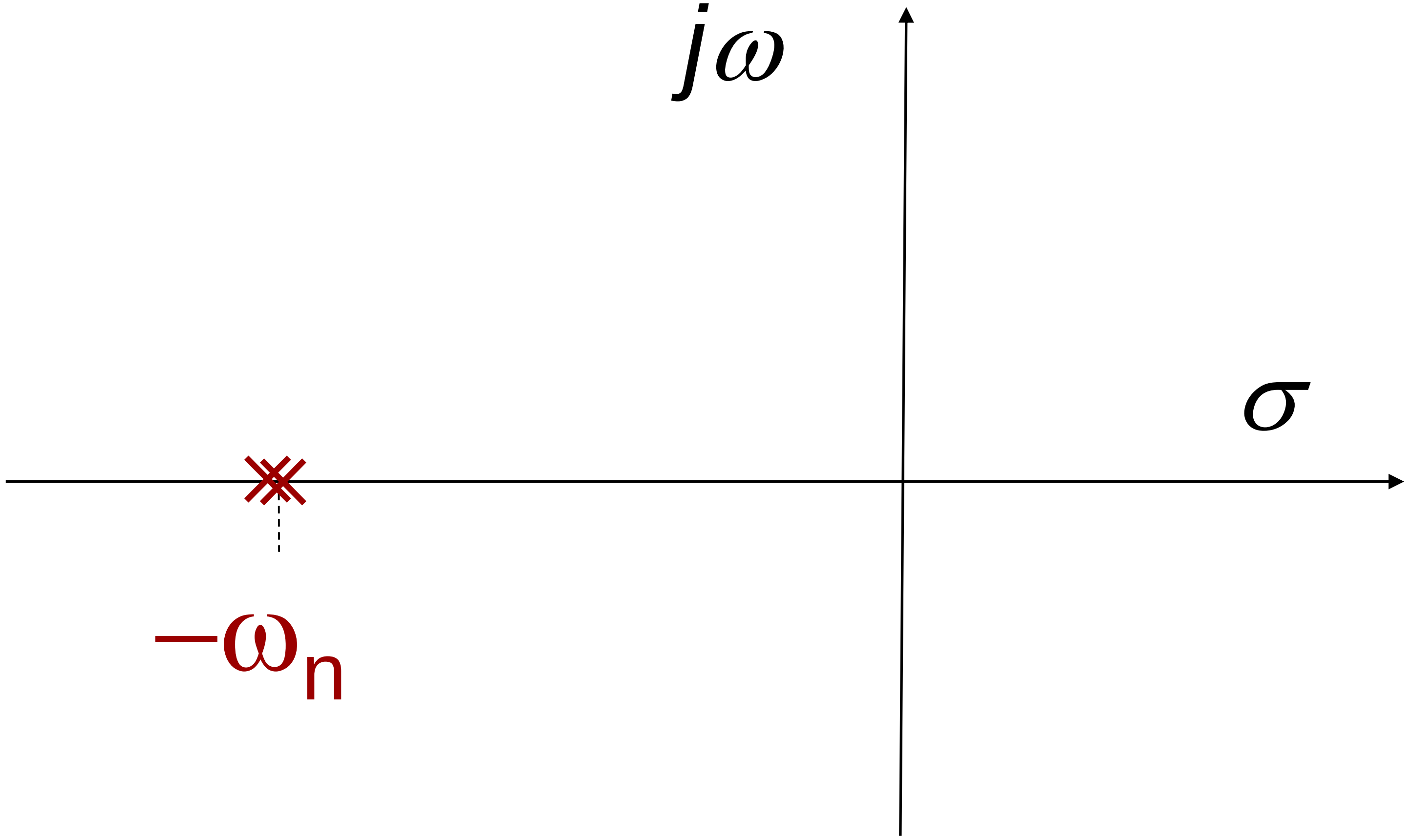


The **RIGHT** most pole with smallest value of σ dominates the response (this was the β term from part 4). Further increasing damping ratio, widens the gap between the two poles and shifts them both further away from the origin.

Conversely, decreasing the damping ratio brings the two poles closer together, until we reach the special case where $\zeta=1$

Critically damped $\zeta = 1$

$$s = -1\omega_n \pm \omega_n \sqrt{2^2 - 1} = -\omega_n$$

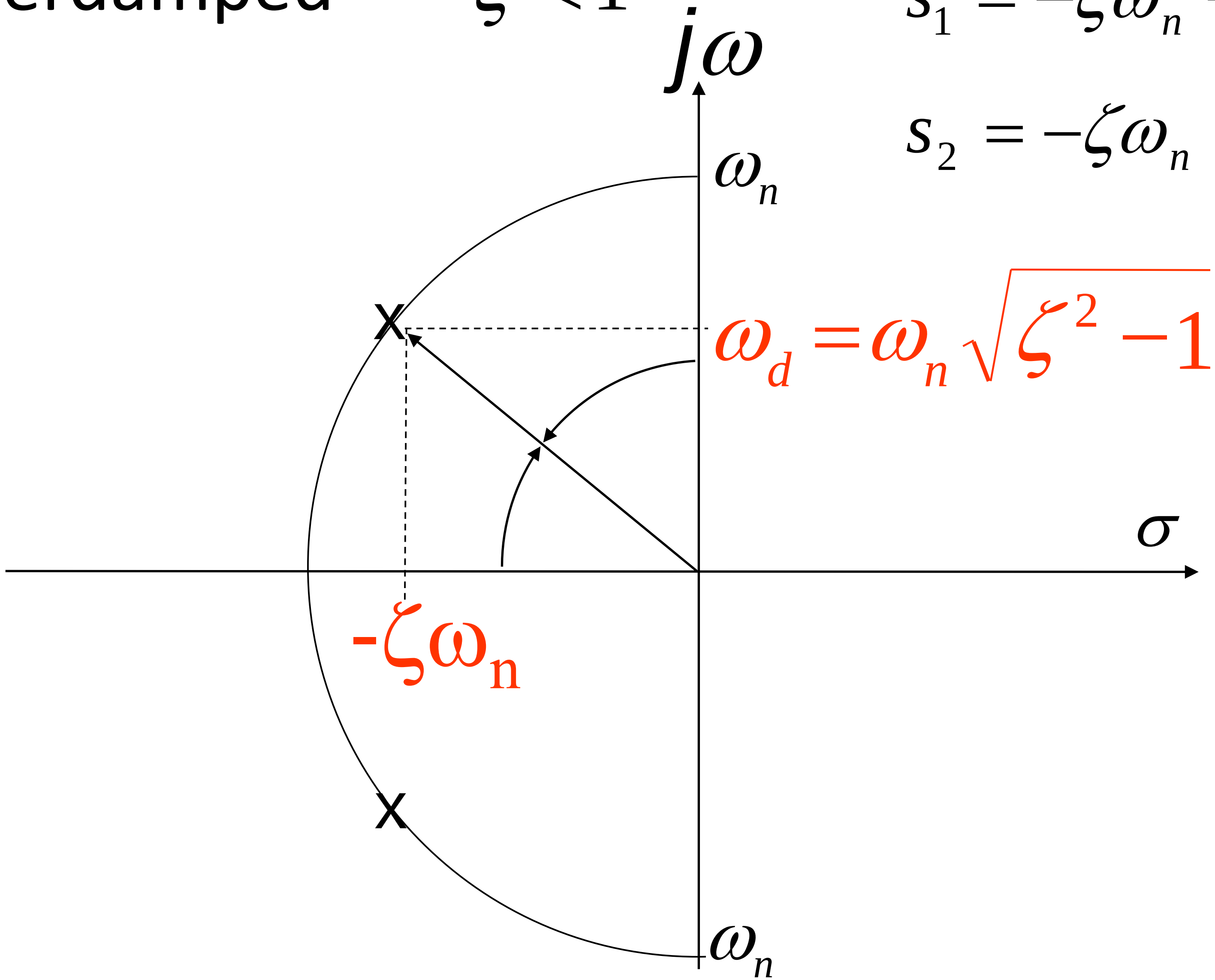


The roots of the equations coincide, and the system is a product of two equal first order lags.

Underdamped $\zeta < 1$

$$s_1 = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}$$

$$s_2 = -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}$$



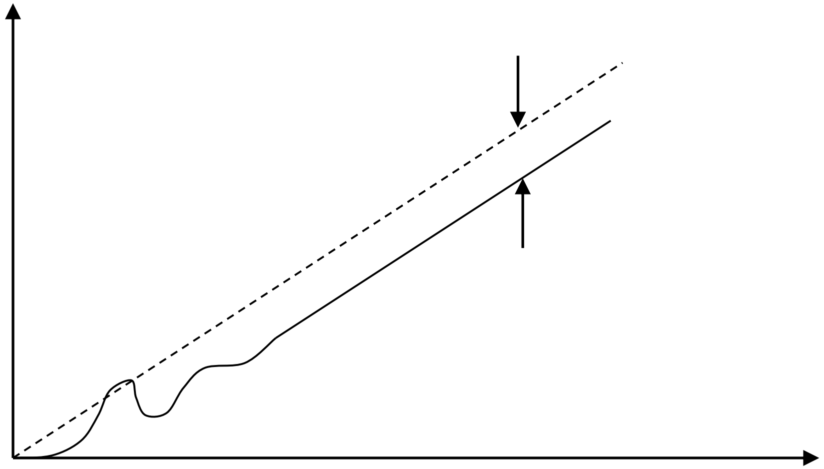
Both roots complex, and form conjugate pair. System is a product of exponential function and oscillatory component.

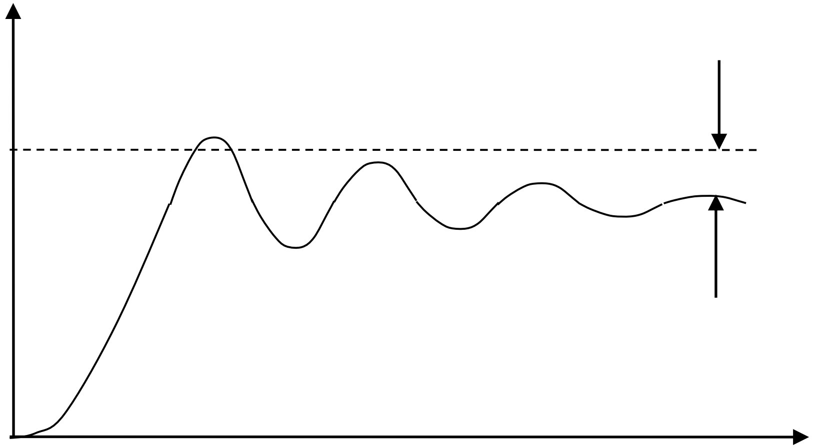
Steady State Performance

Previously we have looked at performance criteria for the transient response of a system: *settling time, peak time, overshoot etc.*

In addition to this it is important to know how accurately the control system tracks the demand once it has settled down, i.e. it is in the *steady state.*

For example, in the case of a ramp or step input:

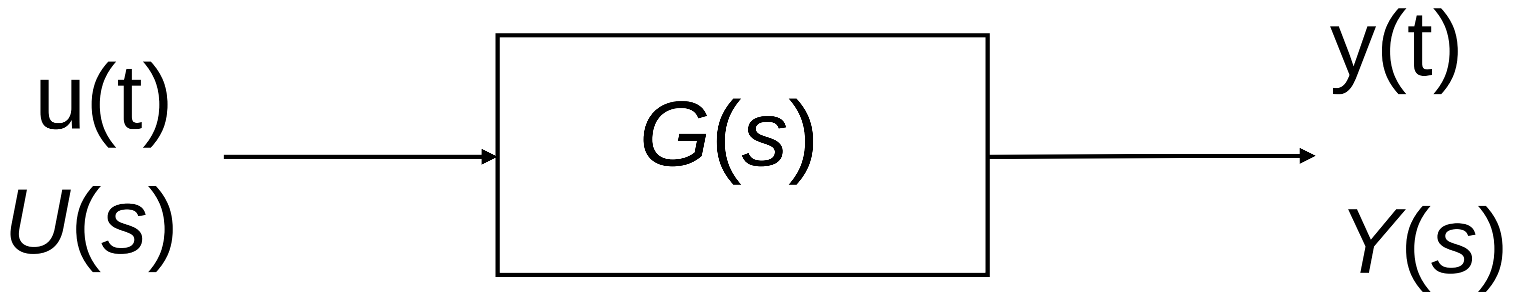




Steady state doesn't imply system is not in motion! – e.g. can be moving at constant velocity, or oscillating.

Final Value Theorem (FVT)

For this we can use the Final Value Theorem, which takes advantage of some handy properties of the Laplace transform



For a constant input, the output will (for some systems) settle down to a steady state value that is a multiple of the input.

Final Value Theorem (FVT)

The final value theorem gives the final value reached:

EXTRA S TERM

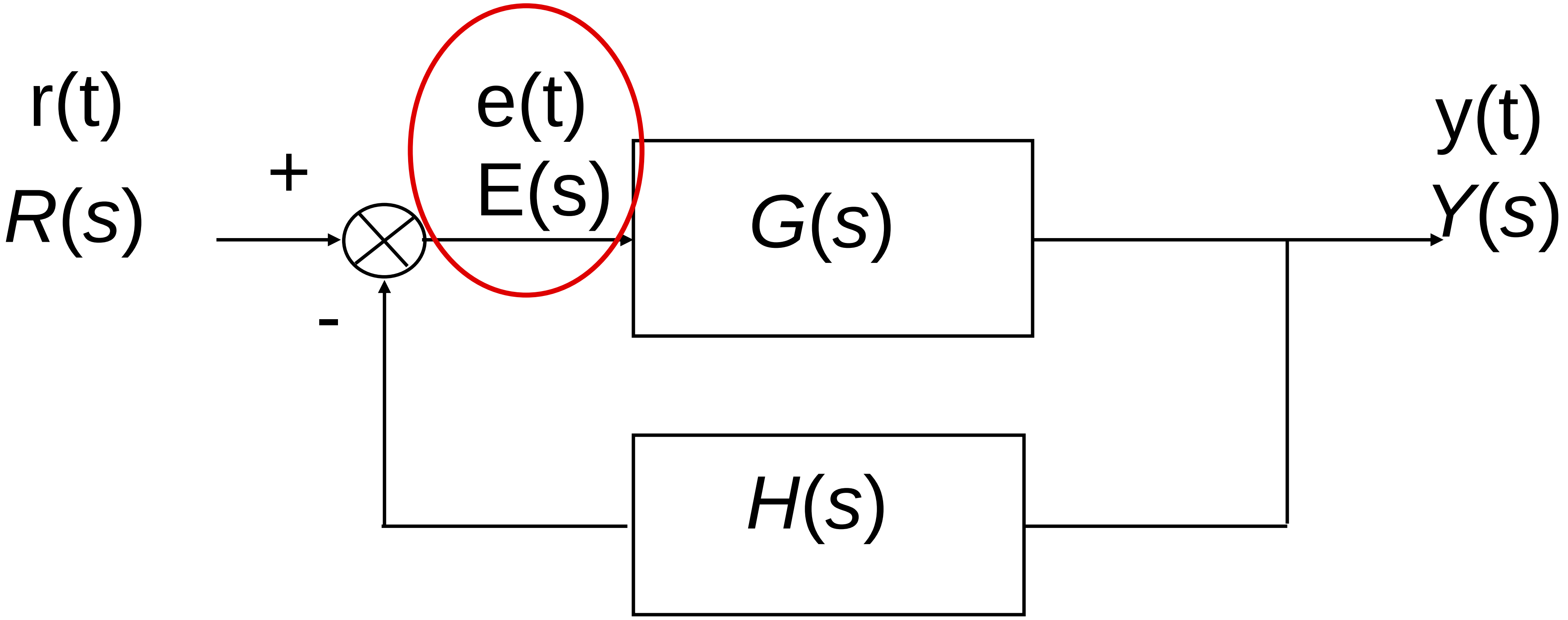
$$\lim_{t \to \infty} y(t) = \lim_{s \to 0} sY(s)$$

So, the final value is found by setting s to zero in the Laplace representation of the output, and multiplying by s

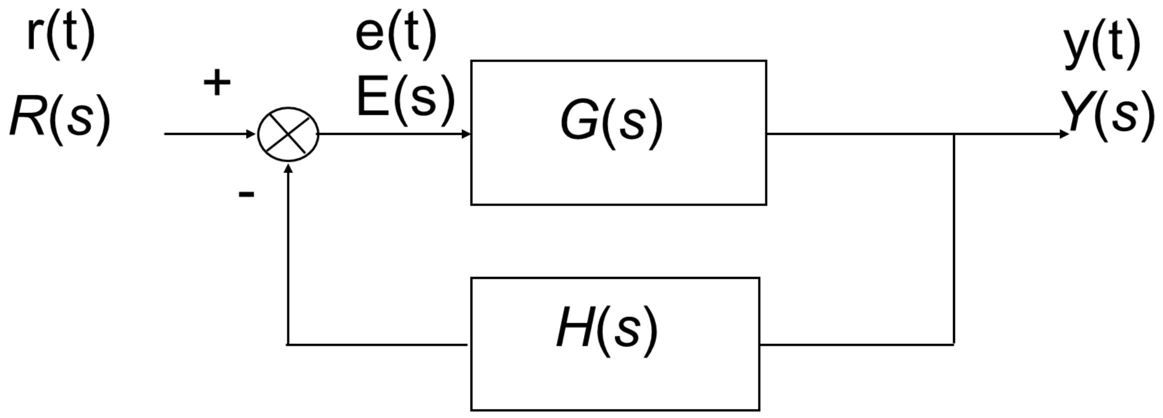
Steady State Error

We are normally more interested in the final value of the *error* rather than the output, as the goal of the controller is to drive the error as close to zero.

Further for some inputs such as a ramp or sinusoidal input, the final output isn't really meaningful.



Steady State Error



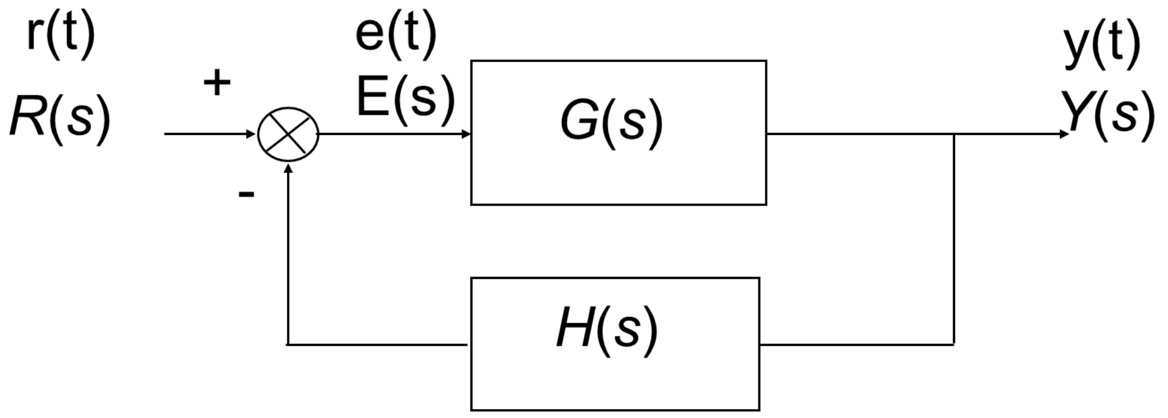
So we first need to write the error signal with respect to the input:

$$E(s) = R(s) - H(s)Y(s) \qquad Y(s) = G(s)E(s)$$

So $\quad E = R - HGE \quad \rightarrow E + HGE = R$

$$E = \frac{R}{1+GH}$$

Steady State Error



Therefore the final value of the error signal e(t) would be:

$$\lim_{t\to\infty} e(t) = \lim_{s\to 0} sE(s)$$

with $$E(s) = \frac{R(s)}{1+G(s)H(s)}$$

so $$\lim_{s\to 0} sE(s) = \lim_{s\to 0} \frac{sR(s)}{1+G(s)H(s)}$$

Steady State Position Error

Using the final value theorem, it is possible to find the steady state error of the control system *and thus judge the appropriateness/success of the controller* without having to calculate the full dynamic response at all.

Remembering back to how long that took with only a second order system, you can imagine how useful this is!

Step input demand

$$R\left(s\right)=\frac{A}{s} \qquad E(s)=\frac{R(s)}{1+G(s)H(s)}$$

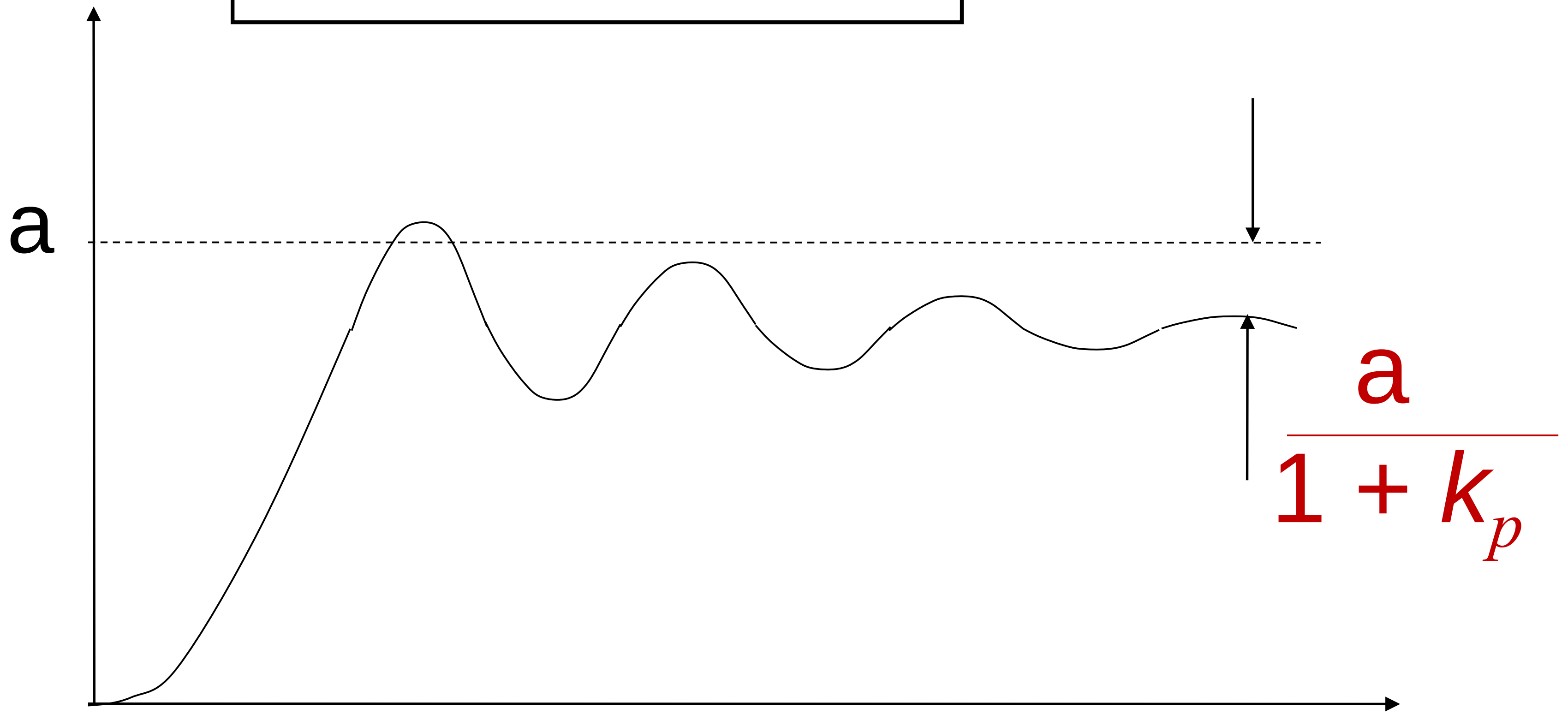
Thus the final error is:

$$\lim_{t\to\infty} e(t)=\lim_{s\to 0}\frac{sR(s)}{1+G(s)H(s)}=\lim_{s\to 0}\frac{s\,{}^{a}\!/_{s}}{1+G(s)H(s)}$$

Step input demand

$$= \lim_{s \to 0} \frac{a}{1+G(s)H(s)} = \frac{a}{1+k_p}$$

where $\boxed{k_p = \lim_{s \to 0} G(s)H(s)}$ Is the **position error constant**



Ramp input demand

$$R(s) = \frac{A}{s^2} \qquad E(s) = \frac{R(s)}{1+G(s)H(s)}$$

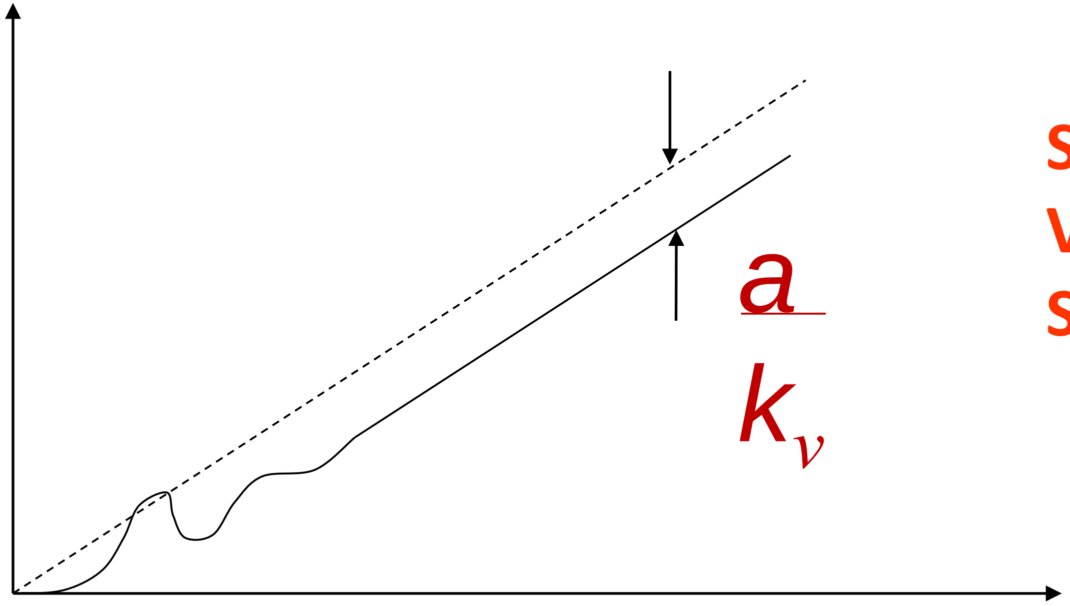
Thus the final error is:

$$\lim_{t\to\infty} e(t) = \lim_{s\to 0} \frac{sR(s)}{1+G(s)H(s)} = \lim_{s\to 0} \frac{s \, {}^{a}\!/_{s^2}}{1+G(s)H(s)}$$

Ramp input demand

$$= \lim_{s \to 0} \frac{a}{s + sG(s)H(s)} = \frac{a}{k_v}$$

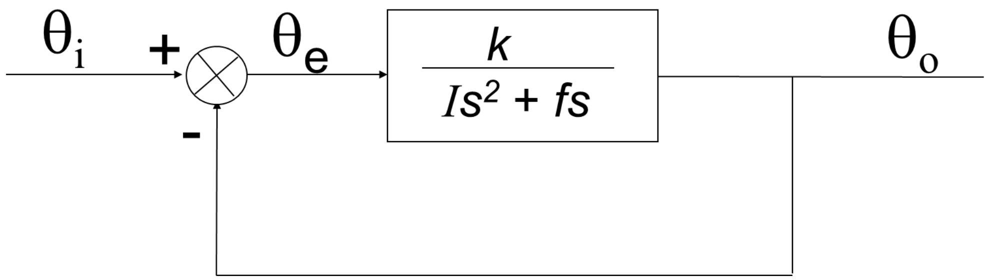
where $\boxed{k_v = \lim_{s \to 0} sG(s)H(s)}$ Is the **velocity error constant**



Steady State Velocity Lag, S.S.V.L.

Once again, lets return to the servo example to see how this system would perform given these inputs

Servo Steady State Error



Step input

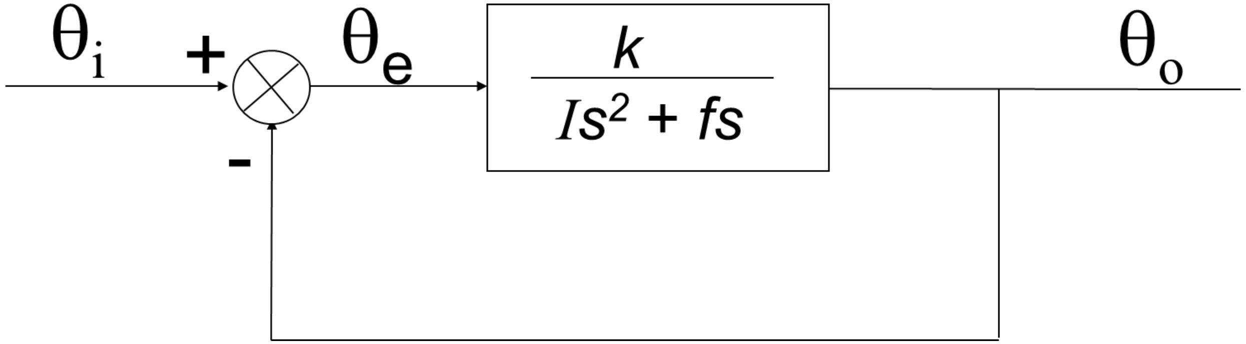
For a unity gain the steady state error is given by $$\frac{a}{1+k_p}$$

$$k_p = \lim_{s \to 0} G(s) = \frac{k}{0.(I.0+f)} = \infty$$

$$\lim_{t \to \infty} \theta_e(t) = \frac{a}{1+k_p} = 0$$

So the position error is zero for all *stable* gains for a step input, *i.e.* The servo matches the demand completely

Servo Steady State Error



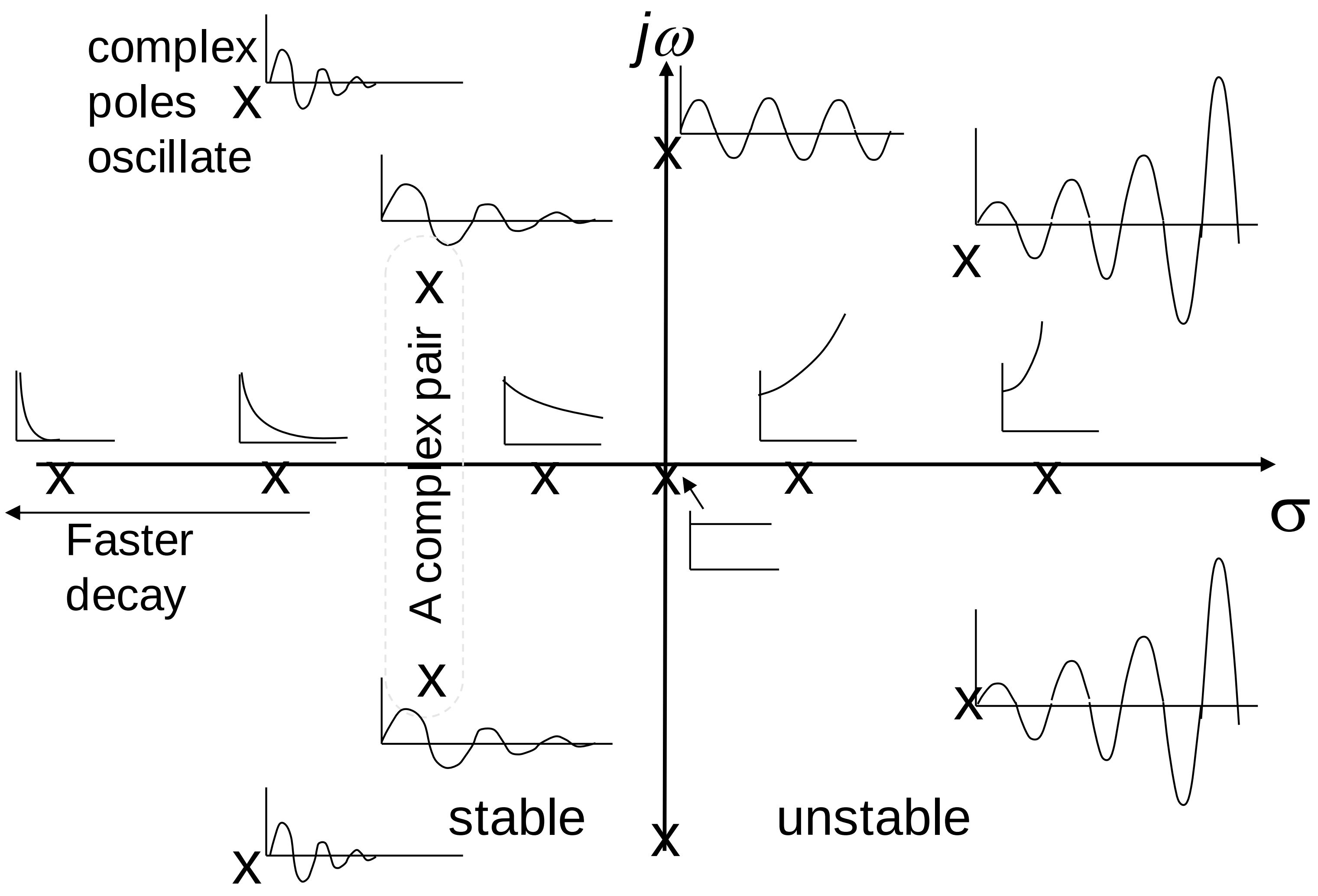
Ramp input

Velocity error constant $$k_v = \lim_{s \to 0} sG(s) = \frac{k}{(Is + f)} = \frac{k}{f}$$

Steady state velocity error $$\lim_{t \to \infty} \theta(t) = \frac{a}{k_v} = \frac{af}{k}$$

This is the lag we saw in lecture 4 when calculating the ramp response directly

Summary


complex
poles
oscillate
$j\omega$
A complex pair
σ
Faster
decay
stable
unstable

Summary

The Final Value Theorem

$$\lim_{t \to \infty} y(t) = \lim_{s \to 0} sY(s)$$

When applied to the Error term, gives the *steady state error without having to calculate the full time domain response*

Thank you for your attention!

Some graphic material used in the course was taken from publicly available online resources that do not contain references to the authors and any restrictions on material reproduction.