B15 Controller Performance

Stephen Duncan
stephen.duncan@eng.ox.ac.uk

Michaelmas Term 2021

Course Aims

The course covers

1. Comparing different representations of dynamic systems
2. A review of closed loop stability
3. Controller design using loop shaping
4. Limitations on achievable controllers
5. A introduction to robustness for single-input, single-output systems

Learning Outcomes

By the end of this course, you should:

1. Be able to describe a system using an ordinary differential equation (ODE), transfer function, impulse response and state space model and be able to switch between these representations
2. Have a deeper understanding of closed loop stability
3. Know how to use loop shaping to design feedback controllers
4. Appreciate the limitations on performance that can be achieved using feedback
5. Understand the importance of designing controllers that make the feedback system robust to uncertainties in the response of the system
6. Know how to use weighting functions and the small-gain theorem to design controllers that provide robustness

Linkages

The aim of this course is to provide a "bridge" between the control theory that you learnt in the 2nd and 3rd years of the course and the more advanced topics in the remainder of the C20 course. The course mainly draws on

1. Introduction to Control Theory (A2)
2. Linear Dynamic Systems (B15)
3. Optimal Control (B15)

You should be familiar with ideas introduced in other courses

1. Ordinary Differential Equations (P1): ODE's, Frequency response, Laplace Transforms, Transfer functions, Poles and Zeros
2. Mathematical Modelling of Physical Systems (P1): Representing physical systems, Block diagrams, Stability
3. Complex Algebra and Fourier Series (P1)
4. Circuit Analysis (P2): Frequency response, Bode diagrams
5. Time-Frequency Analysis (A1): Fourier transforms, Parseval's theorem
6. Linear Algebra (A1): Eigenvalue decompositions

The course leads into

1. Robust Control (C20)

Books

This course sits between introductory courses and the more advanced courses. As a result, there are no books that cover the whole of the course. Most of the introductory material (mainly covered in the first two lectures) are covered in

1. *Feedback Control of Dynamic Systems*, Franklin, G.F., Powell, J.D. and Emami-Naeini, A., (2008) Pearson Education, 5th edition, ISBN 978-0135071816
2. *Modern Control Systems*, Dorf, R.C. and Bishop, R.H. (2011) Pearson Education, 12th edition, ISBN 978-0131383104
3. *Feedback Systems: An Introduction for Scientists and Engineers*, Astrom, K.J. and Murray, R.M. (2008) Princeton University Press, ISBN 978-0691135762

The more advanced material is covered in

1. *Multivariable Feedback Control: Analysis and Design*, Skogestad, S. and Postlethwaite, I. (2005) Wiley, 2nd edition, ISBN 978-0470011683
2. *Linear Optimal Control: H_2 and H_∞ Methods*, Burl, J.B. (1998), Addison-Wesley, ISBN 978-0201808681
3. *Control Theory: Multivariable and Nonlinear Methods*, Glad, T. and Ljung, L. (2000) Taylor and Francis, ISBN 978-0748408788
4. *Control System Design*, Goodwin, G.C., Graebe, S.F. and Salgado, M.E. (2000) Prentice-Hall, ISBN 978-0139586538
5. *Linear Robust Control*, Green, M. and Limebeer, D.J.N. (2012) Dover Publications, ISBN 978-0486488363

However, although these books cover most of the material in this course, they focus on multivariable or multi-input, multi-output (MIMO) systems, which are significantly more complex than the single-input, single-output (SISO) systems that we consider.

Lecture 1

Model Descriptions

Topics

1. Aim of control design
2. Different ways of representing systems
3. Deriving the time response of the system for these representations

Learning Outcomes

1. Understand how to represent a system using ordinary differential equations (ODE's), transfer functions, impulse responses and state space models
2. Appreciate the relationships between the different representations
3. Derive the time response of the systems for each of these representations

1.1 Aims of Control

The problem of designing a control system can be summarised as

Given a system $\mathcal{G}$, with measured signal y, determine feasible input u, so that a controlled variable z as closely as possible follows a reference signal (or setpoint) r, despite the influence of disturbances w, measurement errors n, and variations in the system.

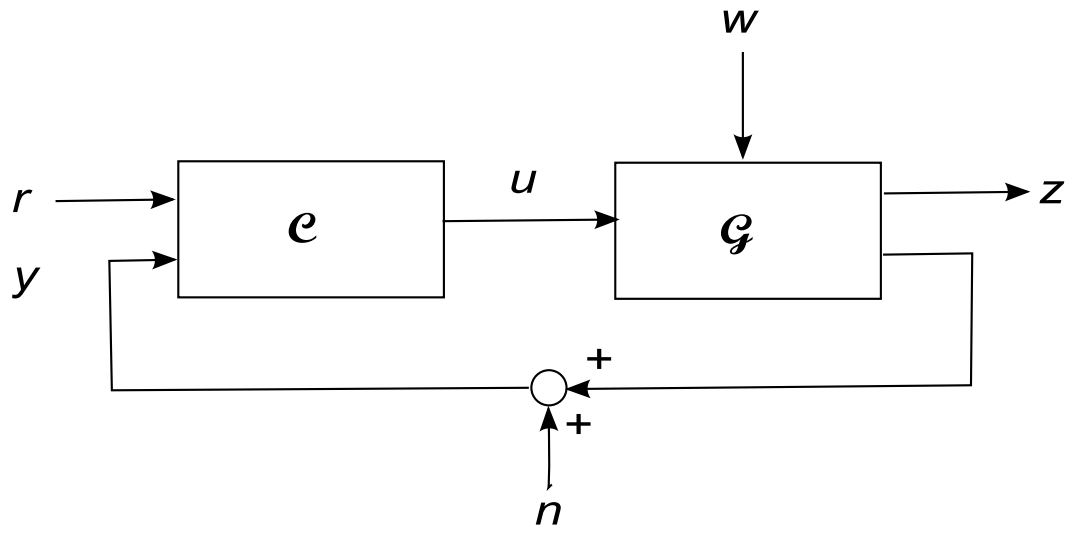

Figure 1.1: Signals in closed loop control system

In general, each of these signals will vary over time, so the signals can be considered as functions of time. For example, we can consider the input signal as a **mapping** from $t \in (-\infty, \infty)$ to the function $u(t)$.

Note. In this course, we will consider signals that are continuous functions of time rather than discrete systems where the signals are a series of samples taken at discrete times, as in the A2 course on Discrete Systems. However, much of what will be discussed can be carried over to discrete systems.

As shown in Figure 1.1, the control problem is typically solved by letting the input signal $u(t)$ be generated automatically as the output from a controller $\mathcal{C}$, which uses the measured signal $y(t)$ and the reference signal $r(t)$ as inputs. Note that in this definition, we have explicitly separated the control variable $z(t)$ from the measured variable $y(t)$. In the majority of control systems, the variable that we want to control is usually the measured variable, before it has been corrupted by the addition of the measurement error. Also, the input to the controller is usually formed from the error signal, which is the difference between the reference signal and the measured signal, so that $e(t) = r(t) - y(t)$. Under these circumstances, the diagram can be redrawn in

the more standard form in Figure 1.2.

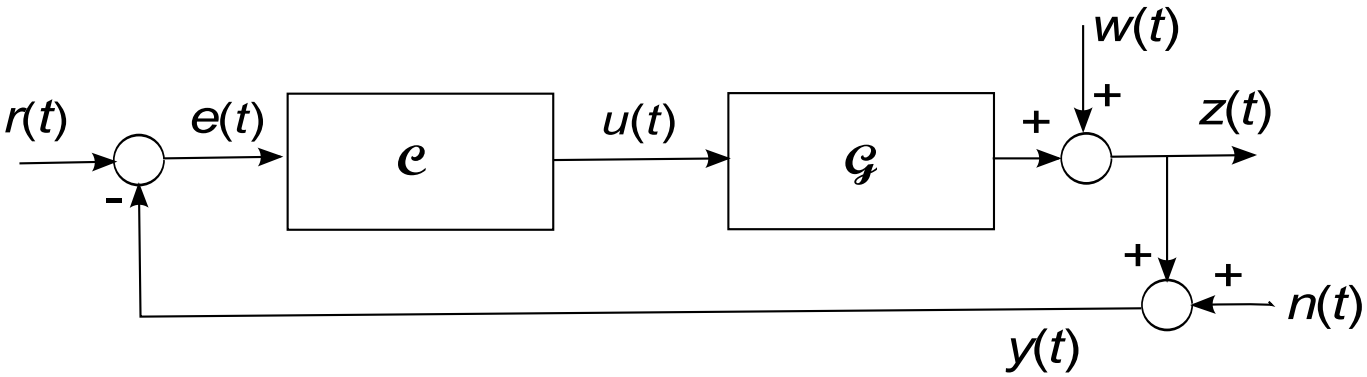

Figure 1.2: Typical arrangement of closed loop control system

For the moment, assume that there is no measurement error, so that $z(t) = y(t)$. The system that we are controlling, $\mathcal{G}$ can be considered as a **mapping** from the input $u(t)$ to the output variable $y(t)$, so that from a mathematical point of view we can write

$$\mathcal{G} : u(t) \mapsto y(t) \tag{1.1.1}$$

or

$$y(t) = \mathcal{G}(u) \tag{1.1.2}$$

Note that $\mathcal{G}$ maps $u(t)$, which is a function of t, to $y(t)$, another function of t.

This is a very general description of a system and in this course, we will restrict our attention to systems that are

Dynamic which means that the output at any time t_1 does not only depend upon the input at this time. If $y(t_1)$ only depends upon $u(t_1)$, then the system is said to be **static**.

Causal where the output does not anticipate the input, which is the case for physical systems. This means that for every point in time t_1, the output $y(t_1)$ depends only on current and past values of the input, $u(t)$ for $\infty < t \leq t_1$, but does not depend on future values of $u(t)$.

Single-input, single-output (SISO) so that at time t_1, all of the signals $u(t_1)$, $y(t_1)$, etc. are scalars (signals that consist of a vector of values are considered in the C20 course on Robust Multivariable Control).

Real so that at any time t_1, the signals are real numbers, $u(t_1) \in \mathbb{R}$, $y(t_1) \in \mathbb{R}$, etc.

Linear so that $\mathcal{G}$ is a linear mapping and the principle of superposition applies

$$\mathcal{G}(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 \mathcal{G}(u_1) + \alpha_2 \mathcal{G}(u_2) \tag{1.1.3}$$

(Nonlinear systems are considered in the C21 course)

Time-invariant systems so that the mapping does not depend upon the absolute value of the time, i.e., if $u(t_1)$ generates an output $y(t_1)$, then $u(t_1 - \tau)$ generates $y(t_1 - \tau)$ for all τ, which means that if the input is shifted in time by a given amount, the corresponding output is also shifted in time by the same amount.

Unconstrained systems This means that there are no limits on the allowable values of the signals. The definition of the the control system says that the controller determines *feasible* inputs. In practice, physical signals have to be limited (constrained control is considered in the C21 course on Model Predictive Control).

You will often see these systems described as dynamic, SISO, LTI (Linear, Time-invariant) systems

1.2 System Descriptions

For dynamic LTI systems, there are a number of different ways to describe the mapping from input to output.

1.2.1 Ordinary Differential Equation

The most basic description is an ordinary differential (ODE) equation with constant coefficients

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y(t) = b_m \frac{\mathrm{d}^m u}{\mathrm{d}t^m} + b_{m-1}\frac{\mathrm{d}^{m-1} y}{\mathrm{d}t^{m-1}} + \cdots + b_1 \frac{\mathrm{d}u}{\mathrm{d}t} + b_0 u(t) \tag{1.2.1}$$

Note that the ODE has been scaled so that the coefficient of the highest differential term on the left hand side is unity.

The **order** of the system is n, which corresponds to the highest differential term on the output (i.e. on the left hand side of the ODE).

Note that if the coefficients in this ODE are functions of time, then the system is still linear, but the underlying system is time varying. In this course, we are considering time invariant systems, so the coefficients are constants.

1.2.2 Transfer Function

Taking Laplace transforms of the ODE in (1.2.1)

$$\begin{aligned}
&\left[s^n Y(s) + s^{(n-1)}y(0) + s^{(n-2)}\dot{y}(0) + \cdots + sy^{(n-2)}(0) + y^{(n-1)}(0)\right] \\
&+ a_{n-1}\left[s^{n-1}Y(s) + s^{(n-2)}y(0) + s^{(n-3)}\dot{y}(0) + \cdots + sy^{(n-3)}(0) + y^{(n-2)}(0)\right] \\
&+ \cdots + a_1\left[sY(s) + y(0)\right] + a_0 Y(s) \\
=&b_m\left[s^m U(s) + s^{(m-1)}u(0) + s^{(m-2)}\dot{u}(0) + \cdots + su^{(m-2)}(0) + u^{(m-1)}(0)\right] \\
&+ b_{m-1}\left[s^{m-1}U(s) + s^{(m-2)}u(0) + s^{(m-3)}\dot{u}(0) + \cdots + su^{(m-3)}(0) + u^{(m-2)}(0)\right] \\
&+ \cdots + b_1\left[sU(s) + u(0)\right] + b_0 U(s) \qquad (1.2.2)
\end{aligned}$$

where $y(0)$, $\dot{y}(0)$, $\ldots$, $y^{(n-1)}(0)$ are the initial conditions on the values of the output and its first $(n-1)$ derivatives at $t = 0$, and similarly for the initial conditions of the input at $t = 0$. Rearranging (1.2.2) gives

$$\begin{aligned}
\left[s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0\right] Y(s) = &\left[b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0\right] U(s) \\
&+ \beta(s) \qquad (1.2.3)
\end{aligned}$$

where $\beta(s)$ is a polynomial of order $(n-1)$ whose coefficients depend upon the initial conditions of the input and output signals.

Define the **transfer function** as

$$G(s) = \frac{b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0} \qquad (1.2.4)$$

then

$$Y(s) = G(s)U(s) + \frac{\beta(s)}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0} \qquad (1.2.5)$$

The first term on the right hand side of (1.2.5) is the Laplace transform of the **particular integral**, while the second term is the Laplace transform of the **complementary function**.

Because we have scaled the ODE so that the first coefficient on the left hand side of (1.2.1) is unity, then the leading coefficient of the denominator polynomial in (1.2.4) is also unity, so the system is **monic**.

If the initial conditions $y(0), \dot{y}(0) \ldots$ and $u(0), \dot{u}(0) \ldots$ are all zero, then all of the coefficients of the polynomial $\beta(s)$ are zero, so (1.2.5) reduces to

$$Y(s) = G(s)U(s) \tag{1.2.6}$$

The transfer function in (1.2.4) is the ratio of a numerator polynomial and a denominator polynomial

$$G(s) = \frac{N(s)}{D(s)} = \frac{b_m s^m + \cdots + b_1 s + b_0}{s^n + \cdots + a_1 s + a_0} \tag{1.2.7}$$

where the order of the numerator polynomial is m and the order of the denominator polynomial is n, which is the order of the system. Because we are considering **causal** systems, then $m \leq n$, so that the system is **proper**; if $m < n$, then we say that the system is **strictly proper**. If $m = n$ (so that the system is proper, but not strictly proper), then we can always write

$$G(s) = \frac{N(s)}{D(s)} = \frac{b_n s^n + b_{n-1} s^{n-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0} = b_n + \frac{\gamma(s)}{D(s)} \tag{1.2.8}$$

where

$$\gamma(s) = N(s) - b_n D(s) \tag{1.2.9}$$

$$= b_n s^n + b_{n-1} s^{n-1} + b_{n-2} s^{n-2} + \cdots + b_1 s + b_0 - b_n s^n - b_n a_{n-1} s^{n-1} - b_n a_{n-2} s^{n-2} - \cdots - b_n a_1 s + b_n a_0 \tag{1.2.10}$$

$$= (b_{n-1} - b_n a_{n-1}) s^{n-1} + (b_{n-2} - b_n a_{n-2}) s^{n-2} + \cdots + (b_1 - b_n a_1) s + (b_0 - b_n a_0) \tag{1.2.11}$$

so that the order of $\gamma(s)$ is $n-1$. This means that when $m = n$, we can rewrite (1.2.6) as

$$Y(s) = G(s)U(s) = b_n U(s) + \frac{\gamma(s)}{D(s)} U(s) \tag{1.2.12}$$

so that the output is the sum of a “straight-through” term, which only scales the input by a static gain b_n, plus the output of a strictly proper system with transfer function $\gamma(s)/D(s)$.

Because the transfer function in (1.2.4) consists of the ratio of two polynomials and by factorising the polynomials, we can write the transfer function in pole-zero form

$$G(s) = \frac{b_m(s - z_1)\dots(s - z_m)}{(s - p_1)\dots(s - p_n)} = \frac{b_m \prod_{k=1}^{m}(s - z_k)}{\prod_{k=1}^{n}(s - p_k)} \tag{1.2.13}$$

where $\{z_k : k = 1, \dots, m\}$ are the **zeros** of the transfer function, which are the roots of the numerator polynomial, while $\{p_k : k = 1, \dots, n\}$ are the **poles**, which are the roots of the denominator polynomial.

Note that the pole-zero form is really just a rearrangement of the transfer function, rather than a separate form of the model, but as we will see, it is a useful representation of the system (and it has its own MATLAB command).

1.2.3 Impulse Response

When the initial conditions are zero, then taking the inverse Laplace transform of (1.2.6)

$$y(t) = \int_{-\infty}^{\infty} g(\tau)u(t - \tau)\mathrm{d}\tau \tag{1.2.14}$$

where $g(t)$ is the inverse Laplace transform of the transfer function $G(s)$. For the inverse Laplace transform to exist, we require that $g(t) = 0$ for $t < 0$, so the limits on the convolution can be restricted, leading to

$$y(t) = \int_{0}^{\infty} g(\tau)u(t - \tau)\mathrm{d}\tau \tag{1.2.15}$$

We call $g(t)$ the **impulse response** of the system, which is the output $y(t)$ obtained when an impulse (delta function) is applied to the input at $t = 0$, so that $u(t) = \delta(0)$. Note also that for this input $U(s) = 1$, so (1.2.6) reduces to $Y(s) = G(s)$. This means that the output of the system is the **convolution** of the input with the impulse response.

Since $G(s)$ is the ratio of two polynomials, we can write any strictly proper transfer function in terms of a partial fraction expansion

$$\begin{aligned} G(s) &= \frac{b_m s^m + b_{m-1}s^{m-1} + \dots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1 s + a_0} \\ &= \frac{R_1}{s - p_1} + \frac{R_2}{s - p_2} + \dots + \frac{R_n}{s - p_n} \end{aligned} \tag{1.2.16}$$

where $\{p_k : k = 1, 2, \ldots, n\}$ are the poles of the transfer function, which equal to the roots of the characteristic equation, and $\{R_k : k = 1, 2, \ldots, n\}$ are the coefficients of the partial fraction expansion. Taking the inverse Laplace transform of (1.2.16) gives the impulse response

$$g(t) = R_1 e^{p_1 t} + R_2 e^{p_2 t} + \cdots + R_n e^{p_n t} \tag{1.2.17}$$

Any complex poles will occur in complex conjugate pairs and the corresponding R_k will also be complex conjugates. If the transfer function is proper (rather than strictly proper) then a static gain will be added to $g(t)$ in order to include the straight-through term.

1.2.4 State Space Model

The state space form can be considered as a rearrangement of the ODE in (1.2.1). To start with, consider the ODE

$$\frac{\mathrm{d}^n y_1}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y_1}{\mathrm{d}t^{n-1}} + a_{n-2}\frac{\mathrm{d}^{n-2} y_1}{\mathrm{d}t^{n-1}} + a_{n-3}\frac{\mathrm{d}^{n-3} y_1}{\mathrm{d}t^{n-3}} + \cdots + a_1\frac{\mathrm{d}y_1}{\mathrm{d}t} + a_0 y_1(t) = u(t) \tag{1.2.18}$$

where there there are no differential terms on the right hand side of the ODE and the static gain is unity (so that $b_0 = 1$). This can be rewritten as

$$\frac{\mathrm{d}^n y_1}{\mathrm{d}t^n} = -a_{n-1}\frac{\mathrm{d}^{n-1} y_1}{\mathrm{d}t^{n-1}} - a_{n-2}\frac{\mathrm{d}^{n-2} y_1}{\mathrm{d}t^{n-1}} - a_{n-3}\frac{\mathrm{d}^{n-3} y_1}{\mathrm{d}t^{n-3}} - \cdots - a_1\frac{\mathrm{d}y_1}{\mathrm{d}t} - a_0 y_1(t) + u(t) \tag{1.2.19}$$

Define a state vector $\mathbf{x}(t) \in \mathbb{R}^n$ as

$$\mathbf{x}(t) = \begin{bmatrix} \frac{\mathrm{d}^{n-1} y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2} y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3} y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} \tag{1.2.20}$$

Note that the state vector has dimension n, which is the same as the order of the ODE in (1.2.1) and the order of the denominator polynomial in the transfer function in (1.2.4).

Using this state vector, (1.2.20) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \frac{\mathrm{d}^{n-1}y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2}y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3}y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} = \begin{bmatrix} -a_{(n-1)} & -a_{(n-2)} & -a_{(n-3)} & \cdots & -a_1 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}^{n-1}y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2}y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3}y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} u(t) \tag{1.2.21}$$

The output $y_1(t)$ can be recovered from the state using the measurement equation

$$y_1(t) = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}^{n-1}y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2}y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3}y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} \tag{1.2.22}$$

This state space model assumed that $b_0 = 1$. If the ODE is

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1}y}{\mathrm{d}t^{n-1}} + a_{n-2}\frac{\mathrm{d}^{n-2}y}{\mathrm{d}t^{n-1}} + a_{n-3}\frac{\mathrm{d}^{n-3}y}{\mathrm{d}t^{n-3}} + \cdots + a_1\frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y(t) = b_0 u(t) \tag{1.2.23}$$

then because the system is linear, the output in this case will be b_0 times the output of the state space model in (1.2.21) and (1.2.22), so that

$$y(t) = b_0 y_1(t) \tag{1.2.24}$$

This means that the equation describing the evolution of the state (1.2.21) remains the same, but the measurement equation for this system is a scaled

version of (1.2.22)

$$y(t) = [0 \;\; 0 \;\; 0 \;\; \cdots \;\; 0 \;\; b_0] \begin{bmatrix} \frac{\mathrm{d}^{n-1} y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2} y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3} y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} \tag{1.2.25}$$

Similarly, if the ODE is

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + a_{n-2}\frac{\mathrm{d}^{n-2} y}{\mathrm{d}t^{n-1}} + a_{n-3}\frac{\mathrm{d}^{n-3} y}{\mathrm{d}t^{n-3}} + \cdots + a_1\frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y(t) = b_1\frac{\mathrm{d}u}{\mathrm{d}t} \tag{1.2.26}$$

then we can still use the same state evolution equation, but the output is now

$$y(t) = b_1\frac{\mathrm{d}y_1}{\mathrm{d}t} \tag{1.2.27}$$

so that the measurement equation becomes

$$y(t) = [0 \;\; 0 \;\; 0 \;\; \cdots \;\; b_1 \;\; 0] \begin{bmatrix} \frac{\mathrm{d}^{n-1} y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2} y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3} y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} \tag{1.2.28}$$

Extending this to the general case of a strictly proper system with $m = n-1$

$$\begin{aligned} &\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + a_{n-2}\frac{\mathrm{d}^{n-2} y}{\mathrm{d}t^{n-1}} + a_{n-3}\frac{\mathrm{d}^{n-3} y}{\mathrm{d}t^{n-3}} + \cdots + a_1\frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y(t) \\ &= b_{n-1}\frac{\mathrm{d}^{n-1} u}{\mathrm{d}t^{n-1}} + b_{n-2}\frac{\mathrm{d}^{n-2} u}{\mathrm{d}t^{n-1}} + b_{n-3}\frac{\mathrm{d}^{n-3} u}{\mathrm{d}t^{n-3}} + \cdots + b_1\frac{\mathrm{d}u}{\mathrm{d}t} + b_0 u(t) \end{aligned} \tag{1.2.29}$$

the state space model is

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \frac{\mathrm{d}^{n-1}y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2}y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3}y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} = \begin{bmatrix} -a_{(n-1)} & -a_{(n-2)} & -a_{(n-3)} & \cdots & -a_1 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}^{n-1}y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2}y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3}y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} u(t) \tag{1.2.30}$$

$$y(t) = [b_{n-1} \quad b_{n-2} \quad b_{n-3} \quad \cdots \quad b_1 \quad b_0] \begin{bmatrix} \frac{\mathrm{d}^{n-1}y_1}{\mathrm{d}t^{n-1}} \\ \frac{\mathrm{d}^{n-2}y_1}{\mathrm{d}t^{n-2}} \\ \frac{\mathrm{d}^{n-3}y_1}{\mathrm{d}t^{n-3}} \\ \vdots \\ \frac{\mathrm{d}y_1}{\mathrm{d}t} \\ y_1(t) \end{bmatrix} \tag{1.2.31}$$

If $m < n-1$, then the terms $b_{m+1}, b_{m+2}, \ldots, b_{n-1}$ in the measurement equation are zero.

This form of state space model is referred to as the **controllable canonical** form and a block diagram representation of the system is shown in Figure 1.3.

The state space model defined by (1.2.30) and (1.2.31) can be written

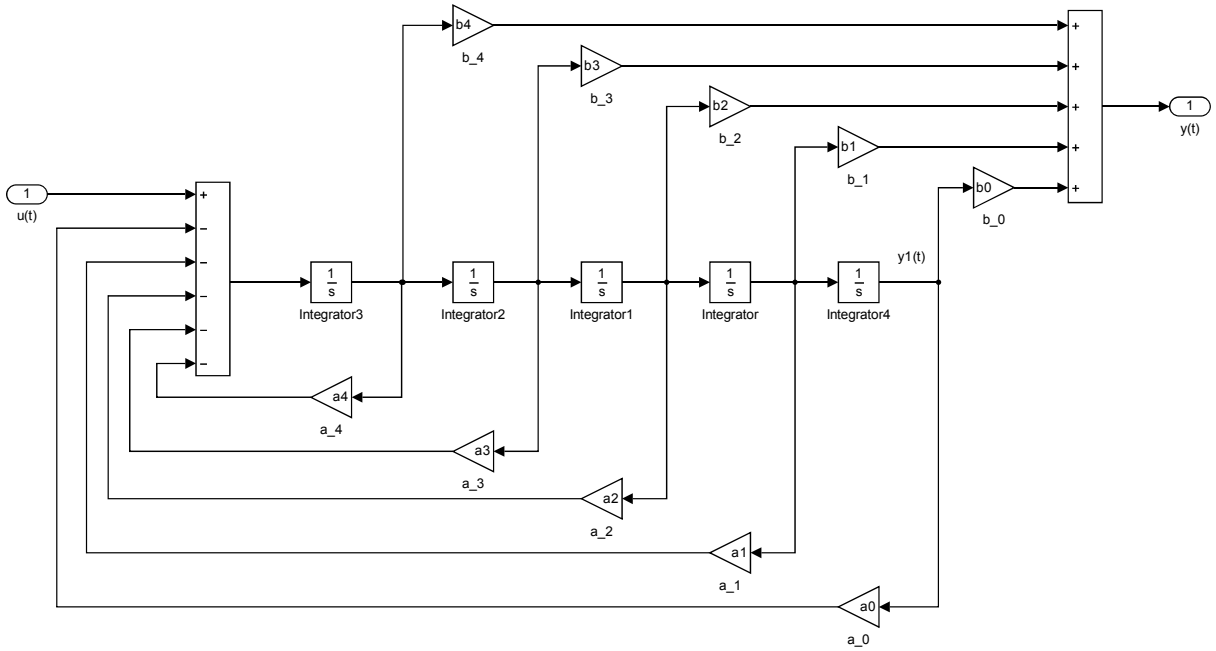

Figure 1.3: Block diagram representation of state space model in controllable canonical form

more compactly by defining $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^n$

$$\mathbf{A} = \begin{bmatrix} -a_{(n-1)} & -a_{(n-2)} & -a_{(n-3)} & \cdots & -a_1 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} b_{n-1} \\ b_{n-2} \\ b_{n-3} \\ \cdots \\ b_1 \\ b_0 \end{bmatrix} \tag{1.2.32}$$

so that

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) &= \mathbf{c}^{\mathsf{T}}\mathbf{x}(t) \end{aligned} \tag{1.2.33}$$

Note 1 you will sometimes see the controllable canonical form with the $-a_{(n-1)}$, $-a_{(n-2)}$, $-a_{(n-3)}$ etc terms across the bottom row of the **A** matrix, rather than the top row. This is because the order of the terms in the state vector has been reversed, so that $y_1(t)$ is the first element in the state vector and $\frac{\mathrm{d}^{n-1}y_1}{\mathrm{d}t^{n-1}}$ is the last element.

Note 2. If $m = n$, so that the system is proper, but not strictly proper, then the system can be separated into a static "straight through" term and a strictly proper dynamic system, as in (1.2.12), so that the state space model

becomes

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)\\ y(t) &= \mathbf{c}^{\mathsf{T}}x(t) + du(t)\end{aligned} \tag{1.2.34}$$

where the scalar d is the static gain and $\mathbf{A}$, $\mathbf{b}$ and $\mathbf{c}$ relate to the strictly proper system. Under these circumstances, $d = b_n$ and using the expression for $\gamma(s)$ in (1.2.11), then

$$\mathbf{c}^{\mathsf{T}} = \begin{bmatrix} b_{n-1} - b_n a_{n-1} & b_{n-2} - b_n a_{n-2} & \cdots & b_1 - b_n a_1 & b_0 - b_n a_0 \end{bmatrix} \tag{1.2.35}$$

This is the form given in the B15 Linear Systems course.

Taking the Laplace transform of (1.2.34) and assuming that the initial conditions of each element of the state vector are zero, then

$$s\mathbf{X}(s) = \mathbf{A}\mathbf{X}(s) + \mathbf{b}U(s) \tag{1.2.36}$$

$$Y(s) = \mathbf{c}^{\mathsf{T}}\mathbf{X}(s) + dU(s) \tag{1.2.37}$$

Rearranging (1.2.36)

$$(s\mathbf{I} - \mathbf{A})\,\mathbf{X}(s) = \mathbf{b}U(s) \tag{1.2.38}$$

and substituting into (1.2.37) gives

$$Y(s) = \left[\mathbf{c}^{\mathsf{T}}\,(s\mathbf{I} - \mathbf{A})^{-1}\,\mathbf{b} + d\right]\,U(s) \tag{1.2.39}$$

Comparing this expression with (1.2.6) allows us to identify that the transfer function associated with a state space model is

$$G(s) = \mathbf{c}^{\mathsf{T}}\,(s\mathbf{I} - \mathbf{A})^{-1}\,\mathbf{b} + d \tag{1.2.40}$$

Note 1. The state space model in (1.2.30) and (1.2.31) can expressed as

$$\left[\begin{array}{c} \dot{\mathbf{x}}(t) \\ y(t) \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{b} \\ \hline \mathbf{c}^{\mathsf{T}} & d \end{array}\right] \left[\begin{array}{c} \mathbf{x}(t) \\ u(t) \end{array}\right] \tag{1.2.41}$$

You will sometimes see this written as

$$\mathbf{G} = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{b} \\ \hline \mathbf{c}^{\mathsf{T}} & d \end{array}\right] \tag{1.2.42}$$

where $\mathbf{G} \in \mathbb{R}^{(n+1)\times(n+1)}$. Be careful: $\mathbf{G}$ is a real matrix of dimensions $(n+1)$ by $(n+1)$, which is not the same as $G(s)$ which a transfer function, although the two terms are related via (1.2.40).

Note 2. Because we are considering single-input, single-output systems, $\mathbf{b}$ and $\mathbf{c}$ are vectors, while d is a scalar. For the multivariable systems with multiple inputs and multiple outputs, these terms are matrices.

1.2.5 State Transformations

The choice of the state vector in (1.2.20) is not unique in the sense that we can choose a different state vector, for which there will be a different state space model, but the mapping from input $u(t)$ to output $y(t)$ will be the same. To see this, introduce a transformation matrix $\mathbf{T} \in \mathbb{R}^{n\times n}$ which transforms the state $\mathbf{x}(t)$ to a new state $\tilde{\mathbf{x}}(t)$

$$\tilde{\mathbf{x}}(t) = \mathbf{T}\mathbf{x}(t) \tag{1.2.43}$$

substituting $\mathbf{x}(t) = \mathbf{T}^{-1}\tilde{\mathbf{x}}(t)$ in (1.2.34)

$$\mathbf{T}^{-1}\dot{\tilde{\mathbf{x}}}(t) = \mathbf{A}\mathbf{T}^{-1}\tilde{\mathbf{x}}(t) + \mathbf{b}u(t) \tag{1.2.44}$$

$$y(t) = \mathbf{c}^{\mathsf{T}}\mathbf{T}^{-1}\tilde{\mathbf{x}}(t) + du(t) \tag{1.2.45}$$

Rearranging 1.2.44 gives

$$\dot{\tilde{\mathbf{x}}}(t) = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\tilde{\mathbf{x}}(t) + \mathbf{T}\mathbf{b}u(t) \tag{1.2.46}$$

so that the transformed state space model can be written as

$$\dot{\tilde{\mathbf{x}}}(t) = \tilde{\mathbf{A}}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{b}}u(t) \tag{1.2.47}$$

$$y(t) = \tilde{\mathbf{c}}^{\mathsf{T}}\tilde{\mathbf{x}}(t) + \tilde{d}u(t) \tag{1.2.48}$$

where

$$\tilde{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1} \qquad \tilde{\mathbf{b}} = \mathbf{T}\mathbf{b} \qquad \tilde{\mathbf{c}}^{\mathsf{T}} = \mathbf{c}^{\mathsf{T}}\mathbf{T}^{-1} \qquad \tilde{d} = d \tag{1.2.49}$$

The transfer function for the transformed state space model is

$$\begin{aligned} G(s) &= \tilde{\mathbf{c}}^{\mathsf{T}} \left(s\mathbf{I} - \tilde{\mathbf{A}}\right)^{-1} \tilde{\mathbf{b}} + \tilde{d} \\ &= \mathbf{c}^{\mathsf{T}}\mathbf{T}^{-1} \left(s\mathbf{I} - \mathbf{TAT}^{-1}\right)^{-1} \mathbf{Tb} + d \\ &= \mathbf{c}^{\mathsf{T}}\mathbf{T}^{-1}\mathbf{T} \left(s\mathbf{I} - \mathbf{A}\right)^{-1} \mathbf{T}^{-1}\mathbf{Tb} + d \\ &= \mathbf{c}^{\mathsf{T}} \left(s\mathbf{I} - \mathbf{A}\right)^{-1} \mathbf{b} + d \end{aligned} \tag{1.2.50}$$

which is the same as the transfer function in (1.2.40), showing that transforming the state does not affect the input-output behaviour of the system.

An important transformation is associated with the eigenvectors of the $\mathbf{A}$ matrix, for the case where all of the eigenvalues are distinct (i.e. there are no repeated eigenvalues) - the case of repeated eigenvalues will be considered below. If $\mathbf{V} \in \mathbb{R}^{n \times n}$ is a matrix whose columns contain the eigenvectors, then

$$\mathbf{AV} = \mathbf{V}\Lambda \tag{1.2.51}$$

where $\Lambda \in \mathbb{R}^{n \times n}$ is a diagonal matrix containing the eigenvalues of $\mathbf{A}$. If we choose $\mathbf{T} = \mathbf{V}^{-1}$, so that $\tilde{\mathbf{A}} = \mathbf{TAT}^{-1} = \Lambda$ then

$$\dot{\tilde{\mathbf{x}}}(t) = \Lambda\, \tilde{\mathbf{x}}(t) + \mathbf{V}^{-1}\mathbf{b}\, u(t) \tag{1.2.52}$$

$$y(t) = \mathbf{c}^{\mathsf{T}}\mathbf{V}\, \tilde{\mathbf{x}}(t) + d\, u(t) \tag{1.2.53}$$

This diagonalises the transformed $\mathbf{A}$ matrix, where the diagonal elements are the eigenvalues, which has the effect of decoupling the dynamics of the transformed states. Note that because the $\mathbf{A}$ matrix is not symmetric, the matrix of eigenvectors is not orthogonal and the eigenvalues can be complex. However, because the matrix is real, any complex eigenvalues will occur in conjugate pairs.

When the state space system has been transformed so that $\tilde{\mathbf{A}}$ is diagonal,

then

$$\begin{bmatrix} \dot{\tilde{x}}_1(t) \\ \dot{\tilde{x}}_2(t) \\ \vdots \\ \dot{\tilde{x}}_n(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \\ \vdots \\ \tilde{x}_n(t) \end{bmatrix} + \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_n \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} \tilde{c}_1 & \tilde{c}_2 & \cdots & \tilde{c}_n \end{bmatrix} \begin{bmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \\ \vdots \\ \tilde{x}_n(t) \end{bmatrix} + \tilde{d}u(t) \qquad (1.2.54)$$

If we assume that all of the initial conditions are zero, so that $\{\tilde{x}_k(0) = 0 : k = 1, 2, \ldots, n\}$, then taking Laplace transforms

$$\begin{bmatrix} s\tilde{X}_1(s) \\ s\tilde{X}_2(s) \\ \vdots \\ s\tilde{X}_n(s) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \tilde{X}_1(s) \\ \tilde{X}_2(s) \\ \vdots \\ \tilde{X}_n(s) \end{bmatrix} + \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_n \end{bmatrix} U(s)$$

$$Y(s) = \begin{bmatrix} \tilde{c}_1 & \tilde{c}_2 & \cdots & \tilde{c}_n \end{bmatrix} \begin{bmatrix} \tilde{X}_1(s) \\ \tilde{X}_2(s) \\ \vdots \\ \tilde{X}_n(s) \end{bmatrix} + \tilde{d}U(s) \qquad (1.2.55)$$

Rearranging the state equation gives

$$\begin{bmatrix} s-\lambda_1 & 0 & \cdots & 0 \\ 0 & s-\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s-\lambda_n \end{bmatrix} \begin{bmatrix} \tilde{X}_1(s) \\ \tilde{X}_2(s) \\ \vdots \\ \tilde{X}_n(s) \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_n \end{bmatrix} U(s) \qquad (1.2.56)$$

so that

$$\begin{bmatrix} \tilde{X}_1(s) \\ \tilde{X}_2(s) \\ \vdots \\ \tilde{X}_n(s) \end{bmatrix} = \begin{bmatrix} s-\lambda_1 & 0 & \cdots & 0 \\ 0 & s-\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s-\lambda_n \end{bmatrix}^{-1} \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_n \end{bmatrix} U(s) \qquad (1.2.57)$$

Substituting into the measurement equation gives

$$Y(s) = \begin{bmatrix} \tilde{c}_1 & \tilde{c}_2 & \cdots & \tilde{c}_n \end{bmatrix} \begin{bmatrix} s-\lambda_1 & 0 & \cdots & 0 \\ 0 & s-\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s-\lambda_n \end{bmatrix}^{-1} \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_n \end{bmatrix} U(s) + \tilde{d}U(s)$$

$$= \left[\frac{\tilde{c}_1\tilde{b}_1}{s-\lambda_1} + \frac{\tilde{c}_2\tilde{b}_2}{s-\lambda_2} + \cdots + \frac{\tilde{c}_n\tilde{b}_n}{s-\lambda_n} + \tilde{d} \right] U(s) \tag{1.2.58}$$

$$= \left[\frac{N(s)}{(s-\lambda_1)(s-\lambda_2)\ldots(s-\lambda_n)} + \tilde{d} \right] U(s) \tag{1.2.59}$$

This allows us to see that the transfer function is

$$G(s) = \frac{N(s)}{(s-\lambda_1)(s-\lambda_2)\ldots(s-\lambda_n)} + \tilde{d} \tag{1.2.60}$$

where $N(s)$ is a polynomial of order $n-1$. Comparing the partial fraction expansion (1.2.58) with the expression in (1.2.16), we can see that $R_k = \tilde{c}_k\tilde{b}_k$.

1.2.6 Example

For the ODE

$$\frac{\mathrm{d}^3y}{\mathrm{d}t^3} + 9\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 26\frac{\mathrm{d}y}{\mathrm{d}t} + 24y(t) = 2\frac{\mathrm{d}u}{\mathrm{d}t} + 2u(t) \tag{1.2.61}$$

the transfer function is

$$G(s) = \frac{2s+2}{s^3+9s^2+26s+24} \tag{1.2.62}$$

Note that the order of the denominator polynomial is 3, while the order of the numerator polynomial is 1, so this system is strictly proper. The transfer function can be written in pole-zero form as

$$G(s) = \frac{2(s+1)}{(s+2)(s+3)(s+4)} \tag{1.2.63}$$

which can be expanded in terms of partial fractions

$$G(s) = -\frac{1}{s+2} + \frac{4}{s+3} - \frac{3}{s+4} \tag{1.2.64}$$

so that applying the inverse Laplace transform, the impulse response is

$$g(t) = -\mathrm{e}^{-2t} + 4\mathrm{e}^{-3t} - 3\mathrm{e}^{-4t} \tag{1.2.65}$$

The controllable canonical state space model of the system is

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} -9 & -26 & -24 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = [0 \;\; 2 \;\; 2] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \tag{1.2.66}$$

The matrix of eigenvectors of the **A** matrix are

$$\mathbf{V} = \begin{bmatrix} 0.8729 & -0.9435 & 0.9684 \\ -0.4364 & 0.3145 & -0.2421 \\ 0.2182 & -0.1048 & 0.0605 \end{bmatrix} \tag{1.2.67}$$

Applying the transformation $\tilde{\mathbf{x}} = \mathbf{T}\mathbf{x}$ where

$$\mathbf{T} = \mathbf{V}^{-1} = \begin{bmatrix} 2.2913 & 16.0390 & 27.4955 \\ 9.5394 & 57.2364 & 76.3151 \\ 8.2614 & 41.3068 & 49.5681 \end{bmatrix} \tag{1.2.68}$$

gives

$$\begin{bmatrix} \dot{\tilde{x}}_1(t) \\ \dot{\tilde{x}}_2(t) \\ \dot{\tilde{x}}_3(t) \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \\ \tilde{x}_3(t) \end{bmatrix} + \begin{bmatrix} 2.291 \\ 9.539 \\ 8.261 \end{bmatrix} u(t)$$

$$y(t) = [-0.4364 \;\; 0.4193 \;\; -0.3631] \begin{bmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \\ \tilde{x}_3(t) \end{bmatrix} \tag{1.2.69}$$

Note that the eigenvalues of the state space model, which are the elements on the diagonal of $\tilde{\mathbf{A}}$, are the same as the poles of the transfer function and that

$$\tilde{c}_1\tilde{b}_1 = -1.0 \quad \tilde{c}_2\tilde{b}_2 = 4.0 \quad \tilde{c}_3\tilde{b}_3 = -3.0 \tag{1.2.70}$$

which are the same as the coefficients of the partial fraction expansion of the transfer function in (1.2.64).

1.2.7 Repeated Eigenvalues

The diagonalisation of the **A** matrix using a transformation matrix consisting of the eigenvectors only applies if there are no repeated eigenvalues, so that the eigenvalues are distinct and the number of linearly independent eigenvectors is the same as the dimension of the A matrix. This is not the case if there are repeated eigenvalues. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix} \tag{1.2.71}$$

The eigenvalues of this matrix are 1, 2 and 3, where 3 is a repeated eigenvalue. If you ask MATLAB to work out the eignevalues and eigenvectors for this matrix using `[V,D] = eig(A);`, it returns

$$\mathbf{D} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1.2.72}$$

$$\mathbf{V} = \begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ 0 & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 & 0 \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{3}} \end{bmatrix} \tag{1.2.73}$$

where the first two columns are identical because there is only one eigenvector associated with the repeated eigenvalue $\lambda = 3$, so the columns of **V** are not linearly independent and $\mathbf{V}^{-1}$ does not exist. (If you do generate **V** using MATLAB, then in fact it can be inverted, but this is due to numerical errors in the `eig` routine which mean that the $(1,2)$ and $(2,2)$ elements are very small - order of 10^{-16} - rather than zero, which makes the condition number of **V** huge).

There is, however, an alternative transformation that reduces **A** to a **Jordan form**, which can be found using `[P,J] = jordan(A);`, so that

$$\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1} \tag{1.2.74}$$

where

$$\mathbf{J} = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1.2.75}$$

and

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \tag{1.2.76}$$

Note that there is a 1 in the first element of the upper diagonal in **J** and that the first, third and fourth columns of **P** are scaled versions of the eigenvectors, but the second column of **P** is independent of the other columns, so that the matrix is non-singular. The columns of **P** are referred to as the generalised eigenvectors.

1.2.8 MATLAB Commands

MATLAB has a number of commands for setting up for different models

- `tf` creates a transfer function form of the model

  ```
  mm_tf = tf(num,den);
  ```

  where `num` and `den` are vectors containing the coefficients of the numerator and denominator polynomials in decreasing order of the powers of s. For example, if

  $$G(s) = \frac{2s+2}{s^3+9s^2+26s+24} \tag{1.2.77}$$

  then

  ```
  num = [ 2    2 ];
  den = [ 1    9    26    24 ];
  ```

- `zpk` creates a "pole-zero-gain" form of the model

```
mm_zpk = zpk(z,p,k);
```

  where z and p are vectors containing the zeros and poles, respectively, and k is the gain, so that if

$$G(s) = \frac{2(s+1)}{(s+2)(s+3)(s+4)} \tag{1.2.78}$$

  then

```
z = [ -1 ];
p = [ -2 -3 -4 ];
k = 2;
```

- `ss` creates a state space model

```
mm_ss = (A,b,c,d);
```

  where `A`, `b`, `c`, `d` are elements of the model, for example

```
A = [ -9 -26 -24 ; ...
        1   0   0 ; ...
        0   1   0 ];
b = [ 1; 0; 0 ];
c = [ 0  2  2 ];
d = 0;
```

In each case, the MATLAB considers the models as **structures**, so for example, the numerator of the transfer function model can be accessed using

```
mm_tf.num;
```

or the **b** vector in the state space model can be accessed using

```
mm_ss.b;
```

The commands can also be used to switch between models, so that for example,

```
mm2_ss = ss(mm_tf);
```

converts the transfer function model `mm_tf` to state space format.

Note. The state model created by MATLAB when using `ss` to convert from another model, has the same structure as the controllable canonical form, but often the elements are scaled to improve the numerical conditioning of the translation; this is important when dealing with large multi-input, multi-output systems.

We can also transform between different state space models using the command `ss2ss`

```
mm3_ss = ss2ss(mm_ss,T);
```

where `T` is a state transition matrix.

Other useful MATLAB commands include

- `residue` returns a partial fraction expansion of the ratio of two polynomials. For

$$G(s) = \frac{2s+2}{s^3+9s^2+26s+24} \tag{1.2.79}$$

Define

```
num = [ 2   2 ];
den = [ 1   9   26   24 ];
```

then

```
[r,p,k] = residue(num,den);
```

returns `r` which is a vector containing the coefficients of the partial fraction expansion, `p` is a vector containing the poles and `k`, the straight-through term for system that are proper, but not strictly proper.

```
r' = [ -3  4  -1 ]
p' = [ -4  -3  -2 ]
k = []
```

Because this transfer function is strictly proper, `k` is an empty vector. This means that we can express the transfer function as

$$G(s) = -\frac{3}{s+4} + \frac{4}{s+3} - \frac{1}{s+2} \tag{1.2.80}$$

Note that MATLAB returns the poles in `p` in decreasing order of magnitude, so the order of the terms in this partial fraction expansion is reversed compared to (1.2.64).

- `impulse` returns the impulse response

```
[z,t] = impulse(mm_tf);
```

returns a vector `y`, which contains the impulse response at the time points in the vector `t`.

- `eig` returns the eigenvalues and eigenvectors of a matrix

```
[V,D] = eig(A);
```

gives `V` the matrix of eigenvectors and `D` a diagonal matrix with the eigenvalues along the diagonal. In a similar manner to `residue`, MATLAB returns the eigenvalues in decreasing order of their magnitudes.

1.3 Time Response

In this section, we compare the different methods for obtaining the solution, i.e. the time response, of the system for a given input.

1.3.1 ODE

For an ODE

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_1\frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y(t) = b_m\frac{\mathrm{d}^m u}{\mathrm{d}t^m} + \cdots + b_1\frac{\mathrm{d}u}{\mathrm{d}t} + b_0 u(t) \tag{1.3.1}$$

the solution consists of the sum of two components,

$$y(t) = y_{\mathrm{PI}}(t) + y_{\mathrm{CF}}(t) \tag{1.3.2}$$

where $y_{\mathrm{CF}}(t)$ is the complementary function, which is the unforced response (i.e. the response in the absence of an input) that depends upon the initial conditions, while $y_{\mathrm{PI}}(t)$ is the response due to the forcing function (i.e. the input), when the initial conditions are zero.

The complementary function satisfies For an ODE

$$\frac{\mathrm{d}^n y_{\mathrm{CF}}}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_1\frac{\mathrm{d}y_{\mathrm{CF}}}{\mathrm{d}t} + a_0 y_{\mathrm{CF}}(t) = 0 \tag{1.3.3}$$

Substituting a trial solution of the form $y_{\mathrm{CF}}(t) = \mathrm{e}^{st}$ leads to the **characteristic equation**

$$s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 = 0 \tag{1.3.4}$$

which can be factorised as

$$\prod_{k=1}^{n}(s - p_k) = (s - p_1)(s - p_2)\ldots(s - p_n) = 0 \tag{1.3.5}$$

where $\{p_k : k = 1, 2. \ldots, n\}$ are the roots of the characteristic equation. If the roots are not repeated, then the solution takes the form

$$y_{\mathrm{CF}}(t) = C_1\mathrm{e}^{p_1 t} + C_2\mathrm{e}^{p_2 t} + \cdots + C_n\mathrm{e}^{p_n t} \tag{1.3.6}$$

where $\{C_k : k = 1, 2, \ldots n\}$ are constants that are determined from the initial conditions. Because $y_{\mathrm{CF}}(t)$ is a real signal, if there are complex poles, they occur in complex conjugate pairs, and the corresponding constants also occur in complex conjugate pairs. If there are repeated roots, then the solution includes terms of the form $(C_k + C_{k+1}t)\mathrm{e}^{p_k t}$.

Note. Although the expression for the complementary function in (1.3.6) has the same form as the expression for the impulse response in (1.2.16), it is important to realise that the coefficients in two expressions are different: for the complementary function, the coefficients C_k are determined from the initial conditions, while for the impulse response, the coefficients R_k are determined from the numerator polynomial of the transfer function.

The particular integral $y_{\mathrm{PI}}(t)$ is the response due to the input signal when the initial conditions are zero, which satisfies

$$\frac{\mathrm{d}^n y_{\mathrm{PI}}}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y_{\mathrm{PI}}}{\mathrm{d}t^{n-1}} + \cdots + a_1\frac{\mathrm{d}y_{\mathrm{PI}}}{\mathrm{d}t} + a_0 y_{\mathrm{PI}}(t) = b_m\frac{\mathrm{d}^m u}{\mathrm{d}t^m} + \cdots + b_1\frac{\mathrm{d}u}{\mathrm{d}t} + b_0 u(t) \tag{1.3.7}$$

and the form of the particular integral depends upon the input signal $u(t)$. For example, if the input is sinusoidal, such as $u(t) = \sin \omega t$, then $y_{\mathrm{PI}}(t) = A\cos \omega t + B\sin \omega t$, where A and B are constants that are determined by substituting this solution into (1.3.7).

1.3.2 Laplace transform

The Laplace transform of the ODE in (1.3.1) has the form

$$Y(s) = G(s)U(s) + \frac{\beta(s)}{s^n + a_{n-1}s^{n-1} \cdots + a_1 s + a_0} \tag{1.3.8}$$

where $G(s)$ is the transfer function

$$G(s) = \frac{b_m s^m + a_{m-1}s^{m-1} \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} \cdots + a_1 s + a_0} \tag{1.3.9}$$

and $\beta(s)$ is a polynomial of order $n - 1$, whose coefficients depend upon the initial conditions. The second term on the right hand side is the Laplace transform of the complementary function and the denominator of this term can be factorised as

$$\begin{aligned} Y_{\mathrm{CF}}(s) &= \frac{\beta(s)}{s^n + a_{n-1}s^{n-1} \cdots + a_1 s + a_0} \\ &= \frac{\beta(s)}{(s-p_1)(s-p_2)\ldots(s-p_n)} \\ &= \frac{C_1}{(s-p_1)} + \frac{C_2}{(s-p_2)} + \cdots + \frac{C_n}{(s-p_n)} \end{aligned} \tag{1.3.10}$$

where $\{p_k : k = 1, 2. \ldots, n\}$ are the poles, i.e. the roots of the denominator polynomial, and the last line is a partial fraction expansion. Applying inverse Laplace transforms

$$y_{\mathrm{CF}}(t) = C_1 \mathrm{e}^{p_1 t} + C_2 \mathrm{e}^{p_2 t} + \cdots + C_n \mathrm{e}^{p_n t} \tag{1.3.11}$$

Comparing (1.3.6) and (1.3.11), the poles of the transfer function are the same as the roots of the characteristic equation, while the constants in (1.3.6) obtained from the initial conditions are the same as the terms in the partial fraction expansion.

Strictly, these expressions are for the case where there are no repeated roots. If there are repeated roots, then there are higher powers of the terms

in the partial fraction expansion, leading to terms of the form $(C_k + C_{k+1}t)e^{p_k t}$ in $y_{CF}(t)$ as in the ODE solution.

The particular integral is the inverse Laplace transform of

$$Y_{PI}(s) = \frac{b_m s^m + a_{m-1}s^{m-1} \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} \cdots + a_1 s + a_0} U(s) \tag{1.3.12}$$

For a given $U(s)$, this can be expressed in terms of partial fractions and the the inverse Laplace transform can be used to obtain $y_{PI}(t)$.

1.3.3 Impulse response

Because the impulse response in (1.2.15) is derived from the transfer function, it implicity assumes that the initial conditions are zero and it ignores the complementary function, so that

$$y_{PI}(t) = \int_0^\infty g(\tau)u(t-\tau)\mathrm{d}\tau \tag{1.3.13}$$

In practice, the impulse response is not often used to find the response of the system to a given input $u(t)$ because of the difficulty in evaluating the convolution integral. However, as we will see in later lectures, the impulse response is an important tool in analysis of control systems, particularly for determining the gain of a system. The impulse response is applicable in discrete time systems, (see A2 course on Discrete Systems) as for these systems, the convolution becomes a sum rather than an integral, which is computationally easier to evaluate.

1.3.4 State space model

For a general state space model

$$\dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{b}\,u(t) \tag{1.3.14}$$

$$y(t) = \mathbf{c}^{\mathsf{T}}\mathbf{x}(t) + d\,u(t) \tag{1.3.15}$$

The state evolution (1.3.14) is a first order (vector) differential equation, which can be solved by writing

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} - \mathbf{A}\,\mathbf{x}(t) = \mathbf{b}\,u(t) \tag{1.3.16}$$

Pre-multiply by the integrating factor $\mathrm{e}^{-\mathbf{A}t}$

$$\mathrm{e}^{-\mathbf{A}t}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} - \mathrm{e}^{-\mathbf{A}t}\mathbf{A}\,\mathbf{x}(t) = \mathrm{e}^{-\mathbf{A}t}\mathbf{b}\,u(t) \tag{1.3.17}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{-\mathbf{A}t}\mathbf{x}(t)\right] = \mathrm{e}^{-\mathbf{A}t}\mathbf{b}\,u(t) \tag{1.3.18}$$

Integrating from 0 to t

$$\mathrm{e}^{-\mathbf{A}t}\mathbf{x}(t) - \mathrm{e}^{-\mathbf{A}0}\mathbf{x}(0) = \int_0^t \mathrm{e}^{-\mathbf{A}\tau}\mathbf{b}\,u(\tau)\mathrm{d}\tau \tag{1.3.19}$$

Since $\mathrm{e}^{-\mathbf{A}0} = \mathbf{I}$, then premultiplying both sides by $\mathrm{e}^{\mathbf{A}t}$ gives

$$\begin{aligned}\mathbf{x}(t) &= \mathrm{e}^{\mathbf{A}t}\mathbf{x}(0) + \mathrm{e}^{\mathbf{A}t}\int_0^t \mathrm{e}^{-\mathbf{A}\tau}\mathbf{b}\,u(\tau)\mathrm{d}\tau \\ &= \mathrm{e}^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t \mathrm{e}^{\mathbf{A}(t-\tau)}\mathbf{b}\,u(\tau)\mathrm{d}\tau \end{aligned} \tag{1.3.20}$$

Given the state at time $t = 0$ and knowing the input signal $u(t)$, (1.3.20) allows us to determine the state $\mathbf{x}(t)$ at all future times.

The output $y(t)$ is obtained by substituting this expression for the state into the measurement equation (1.3.15)

$$\begin{aligned}y(t) &= \mathbf{c}^{\mathsf{T}}\mathrm{e}^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t \mathbf{c}^{\mathsf{T}}\mathrm{e}^{\mathbf{A}(t-\tau)}\mathbf{b}\,u(\tau)\mathrm{d}\tau + d\,u(t) \\ &= \mathbf{c}^{\mathsf{T}}\mathrm{e}^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t \left[\mathbf{c}^{\mathsf{T}}\mathrm{e}^{\mathbf{A}(t-\tau)}\mathbf{b} + d\,\delta(t-\tau)\right]\,u(\tau)\mathrm{d}\tau \end{aligned} \tag{1.3.21}$$

The first term on the right hand side, which depends upon the initial conditions, is the complementary function, while the second term, which depends upon the input signal $u(t)$, is the particular integral.

Assume that the state has been transformed using the matrix of eigenvectors $\mathbf{V}$, so that $\tilde{\mathbf{x}}(t) = \mathbf{V}^{-1}\mathbf{x}(t)$, then the $\tilde{\mathbf{A}}$ matrix is diagonal, with the diagonal elements being the eigenvalues of the $\mathbf{A}$ matrix, so that

$$\dot{\tilde{\mathbf{x}}}(t) = \Lambda\,\tilde{\mathbf{x}}(t) + \tilde{\mathbf{b}}\,u(t) \tag{1.3.22}$$

$$y(t) = \tilde{\mathbf{c}}^{\mathsf{T}}\,\tilde{\mathbf{x}}(t) + \tilde{d}\,u(t) \tag{1.3.23}$$

where $\tilde{\mathbf{c}}^{\mathsf{T}} = \mathbf{c}^{\mathsf{T}}\mathbf{V}$, $\tilde{\mathbf{b}} = \mathbf{V}^{-1}b$ and $\tilde{d} = d$. The complementary function in (1.3.21) can be written as

$$\begin{aligned}y_{\mathrm{CF}}(t) &= \tilde{\mathbf{c}}^{\mathsf{T}}\mathrm{e}^{\Lambda t}\tilde{\mathbf{x}}(0) \\ &= \tilde{c}_1\tilde{x}_1(0)\,\mathrm{e}^{\lambda_1 t} + \tilde{c}_2\tilde{x}_2(0)\,\mathrm{e}^{\lambda_2 t} + \cdots + \tilde{c}_n\tilde{x}_n(0)\,\mathrm{e}^{\lambda_n t} \end{aligned} \tag{1.3.24}$$

Comparing this with the expressions for the complementary functions in (1.3.6) and (1.3.11), we can see that $\{\lambda_k = p_k : k = 1, 2, \ldots, n\}$, so that the eigenvalues of the **A** matrix are the same as the roots of the characteristic equation and the poles of the transfer function. Also, the terms $\{\tilde{c}_k \tilde{x}_k(0) : k = 1, 2, \ldots, n\}$ are the coefficients C_k.

For the particular integral term

$$y_{\mathrm{PI}}(t) = \int_0^t \left[\mathbf{c}^{\mathsf{T}} \mathrm{e}^{\mathbf{A}(t-\tau)}\mathbf{b} + d\,\delta(t-\tau)\right] u(\tau)\mathrm{d}\tau \tag{1.3.25}$$

with the expression in (1.3.13), we can see that the impulse response is

$$g(t) = \mathbf{c}^{\mathsf{T}} \mathrm{e}^{\mathbf{A}t}\mathbf{b} + d\,\delta(t) \tag{1.3.26}$$

Given that the Laplace transform of $\mathrm{e}^{\mathbf{A}t}$ is $(s\mathbf{I} - \mathbf{A})^{-1}$ (note the analogy with the scalar case, where the Laplace of e^{at} is $1/(s-a)$), then

$$G(s) = \mathbf{c}^{\mathsf{T}} (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} + d \tag{1.3.27}$$

which agrees with the expression in (1.2.40).

Summary

The key point is that

The *roots* of the characteristic equation of the ODE

are the same as

the *poles* of the transfer function

are the same as

the *eigenvalues* of the A matrix in the state space model

Lecture 2

Review of Stability

Topics

1. Definitions of Stability
2. Derivation of Nyquist stability criterion for closed loop systems
3. Using the Nyquist stability criterion when there are open loop poles and/or zeros on the imaginary axis
4. Combining representations of systems
5. Internal stability

Learning Outcomes

1. Understand that a system will be stable if its poles lie in the open right half plane
2. Understand the derivation of the Nyquist criterion for closed loop stability
3. Understand how the Nyquist criterion is modified if there are open loop poles and/or zeros on the imaginary axis
4. Understand the idea of the internal stability in a feedback loop
5. Know how to derive the state space models of systems connected in series, parallel and feedback

2.1 Stability

2.1.1 Unforced Response

The unforced response is the response when there is no input (forcing function), so it is the response due to the initial conditions, which is described by the **complementary function**. We know that the time response of the complementary function takes the form

$$y_{\mathrm{CF}}(t) = C_1 e^{p_1 t} + C_2 e^{p_2 t} + \cdots + C_n e^{p_n t} \tag{2.1.1}$$

where $\{p_k : k = 1, 2, \ldots, n\}$ are the roots of the characteristic equation, which are the same as the poles of the transfer function and the eigenvalues of the **A** matrix. If none of the roots of the characteristic equation are repeated, so that the p_k are distinct, then in order for $y_{\mathrm{CF}}(t)$ to remain finite as $t \to \infty$, we require that the real part of $p_k \leq 0$. If the root is not repeated and **on** the the imaginary axis (so that the real part of $p_k = 0$), then the magnitude of $y_{\mathrm{CF}}(t)$ remains constant, although the response will oscillate if there is a pair of complex conjugate poles on the imaginary axis. However, if there are repeated roots on the imaginary axis, for example, if there are two roots at zeros, for example, $p_k = 0$ and $p_{k+1} = 0$, then $y_{\mathrm{CF}}(t)$ will contain the term $(C_k + C_{k+1} t)$ which is not finite as $t \to \infty$.

Usually, we require the more restrictive condition that $y_{\mathrm{CF}}(t) \to 0$ at $t \to \infty$, so that the complementary function decays to zero, rather than simply remaining finite. In this case, we require that the real part of the root/pole/eigenvalue must be strictly less than zero so that $\{\mathrm{Re}(p_k) < 0 : k = 1, 2, \ldots, n\}$, which will occur if the root/pole/eigenvalue lies in the **open left half plane**.

Note. The **open** left half of the complex plane (OLHP) is the values of s for which $\mathrm{Re}(s) < 0$, which does not include values of s lie on the imaginary axis. The **closed** left half of the complex plane (CLHP) includes the values of s **on** the axis, i.e. all values of s for which $\mathrm{Re}(s) \leq 0$.

2.1.2 Forced Response

The forced response, which is the response due to the input $u(t)$, is the **particular integral** and we define the stability of the forced response in terms of **bounded-input, bounded-output (BIBO) stability**. A **bounded** input is one for which there is a finite constant $N \in \mathbb{R}$ such that

$$|u(t)| \leq N \quad \text{for all } t \in (-\infty, \infty) \tag{2.1.2}$$

A system is bounded-input, bounded-output stable if it produces a bounded signal at the output when any bounded signal is applied at the input.

One way of checking this is to use the impulse response. The particular integral can be expressed as

$$y_{\text{PI}}(t) = \int_0^\infty g(\tau)u(t-\tau)\mathrm{d}\tau \tag{2.1.3}$$

and using (2.1.2)

$$\begin{aligned} |y_{\text{PI}}(t)| &\leq \int_0^\infty |g(\tau)|\,|u(t-\tau)|\,\mathrm{d}\tau \\ &\leq N\int_0^\infty |g(\tau)|\,\mathrm{d}\tau \end{aligned} \tag{2.1.4}$$

Hence, the output will be bounded if

$$\int_0^\infty |g(\tau)|\,\mathrm{d}\tau < \infty \tag{2.1.5}$$

An alternative way of determining BIBO stability is to consider the transfer function form of the system and it is easiest to see this through an example. suppose that we apply a step input to a system with a transfer function $G(s)$, then

$$\begin{aligned} Y_{\mathbf{PI}}(s) &= G(s)U(s) \\ &= \frac{b_m s^m + a_{m-1}s^{m-1}\cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1}\cdots + a_1 s + a_0}\frac{1}{s} \\ &= \frac{b_m s^m + a_{m-1}s^{m-1}\cdots + b_1 s + b_0}{(s-p_1)(s-p_2)\ldots(s-p_{n-1})(s-p_n)}\frac{1}{s} \end{aligned} \tag{2.1.6}$$

If none of the poles of the transfer function are at $s = 0$, then we can write (2.1.6) as a partial fraction expansion

$$Y_{\mathbf{PI}}(s) = \frac{D_1}{s-p_1} + \frac{D_2}{s-p_2} + \cdots + \frac{D_{n-1}}{s-p_{n-1}} + \frac{D_n}{s-p_n} + \frac{D_{n+1}}{s} \tag{2.1.7}$$

where $\{D_k : k = 1, 2, \ldots, n, n+1\}$ are the coefficients of the partial fraction expansion. Taking the inverse Laplace transform,

$$y_{\mathbf{PI}}(t) = D_1 e^{p_1 t} + D_2 e^{p_2 t} + \cdots + D_{n-1} e^{p_{n-1} t} + D_n e^{p_n t} + D_{n+1} \tag{2.1.8}$$

which will be bounded provided that $\{\mathrm{Re}(p_k) \leq 0 : k = 1, 2, \ldots, n\}$. However, if one (or more) of the poles of the transfer function are at $s = 0$, so that

$$\begin{aligned} Y_{\mathbf{PI}}(s) &= \frac{b_m s^m + a_{m-1} s^{m-1} \cdots + b_1 s + b_0}{(s - p_1)(s - p_2) \ldots (s - p_{n-1}) s} \frac{1}{s} \\ &= \frac{D_1}{s - p_1} + \frac{D_2}{s - p_2} + \cdots + \frac{D_n}{s} + \frac{D_{n+1}}{s^2} \end{aligned} \tag{2.1.9}$$

(in this example, we have assumed that $p_n = 0$) and

$$y_{\mathbf{PI}}(t) = D_1 e^{p_1 t} + D_2 e^{p_2 t} + \cdots + D_{n-1} e^{p_{n-1} t} + D_n + D_{n+1} t \tag{2.1.10}$$

which is unbounded due to the presence of the term $D_{n+1} t$.

In general, if there is a pole on the imaginary axis, then it is always possible to find an input that will produce an output that is unbounded. Hence to ensure BIBO stability, we require that all of the poles of the transfer function lie in the open left half plane (OLHP), so that $\{\mathrm{Re}(p_k) < 0 : k = 1, 2, \ldots, n\}$.

2.1.3 Minimal Realisation of a State Space Model

When testing the stability of a system from the location of the eigenvalues of the $\mathbf{A}$ matrix in a state space model, we need to be careful when there are uncontrollable and/or unobservable states in the model. For a single-input, single-output model, one way of detecting controllable states is to transform the state space model using $\mathbf{T} = \mathbf{V}^{-1}$, where $\mathbf{V}$ is the matrix containing the eigenvectors of $\mathbf{A}$, so that

$$\dot{\tilde{\mathbf{x}}}(t) = \Lambda \tilde{\mathbf{x}}(t) + \tilde{\mathbf{b}} u(t) \tag{2.1.11}$$

$$y(t) = \tilde{\mathbf{c}}^{\mathsf{T}} \tilde{\mathbf{x}}(t) + \tilde{d} u(t) \tag{2.1.12}$$

where $\tilde{\mathbf{b}} = \mathbf{V}^{-1}\mathbf{b}$ and $\tilde{\mathbf{c}}^{\mathsf{T}} = \mathbf{c}^{\mathsf{T}}\mathbf{V}$, so that

$$\begin{bmatrix} \dot{\tilde{x}}_1(t) \\ \dot{\tilde{x}}_2(t) \\ \vdots \\ \dot{\tilde{x}}_n(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \\ \vdots \\ \tilde{x}_n(t) \end{bmatrix} + \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_n \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} \tilde{c}_1 & \tilde{c}_2 & \cdots & \tilde{c}_n \end{bmatrix} \begin{bmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \\ \vdots \\ \tilde{x}_n(t) \end{bmatrix} + \tilde{d}u(t) \tag{2.1.13}$$

If an element of $\tilde{\mathbf{b}}$ is zero, then the input $u(t)$ does not affect the corresponding state, so that this state is **uncontrollable**. Similarly, if an elements of $\tilde{\mathbf{c}}$ is zero, then the corresponding state does not contribute to the output, so this state is **unobservable**. Because the uncontrollable states are not affected by the input and the unobservable states do not contribute to the output, then the uncontrollable and unobservable states do not contribute to the overall input-output behaviour of the system, i.e. the mapping from $u(t)$ to $y(t)$. This means that in principle, the eigenvalues associated with uncontrollable and/or unobservable states could lie in the right half plane, so that these states are unstable, but because these states do not affect the mapping from input to output, the input-output mapping will still be stable. To avoid this problem, we assume that the uncontrollable and unobservable states have been eliminated from the state vector $\tilde{\mathbf{x}}(t)$ which can be done without affecting the input-output mapping. A system where the uncontrollable and unobservable states have been removed is called a **minimal realisation**.

This discussion is based on single-input, single-output systems and it also assumes that the eigenvalues of the $\mathbf{A}$ matrix are distinct (i.e. the eigenvalues are not repeated); determining the uncontrollable and unobservable states in multi-input, multi-output systems is less straightforward and requires checking the rank of the controllability and observability grammians - see lecture notes from B15 Linear Systems course.

2.2 Nyquist Stability Criterion

Ensuring that the controller is designed so that the closed loop system is stable is an essential pre-requisite; there is no point in designing a controller that will (theoretically) give make the controller variable match the reference signal if the system is unstable!

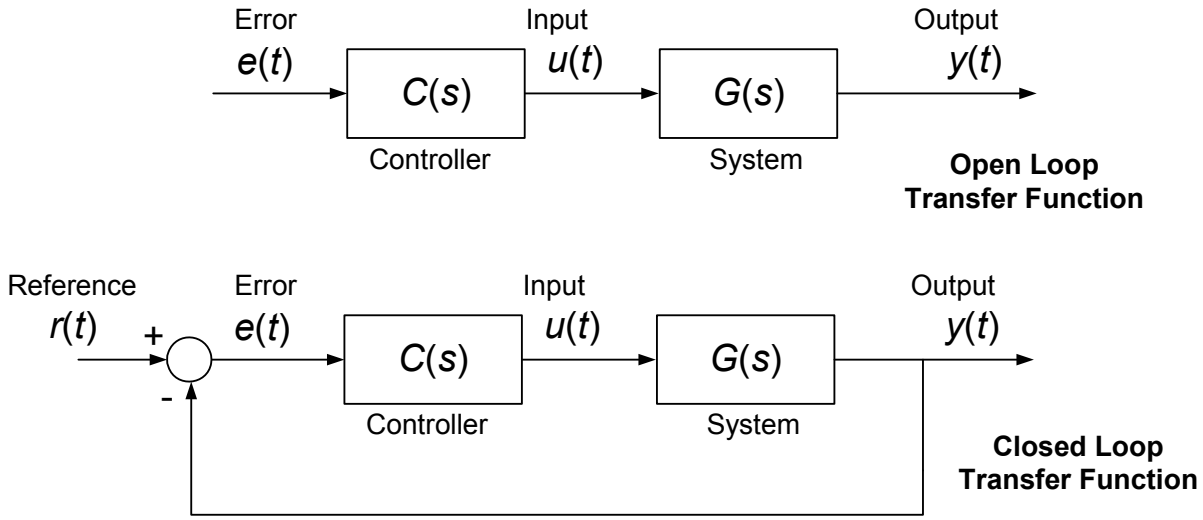

Figure 2.1: Block diagrams of open loop and closed loop systems

In order for a closed loop system with a feedback controller to be stable, we need to design the transfer function of the controller $C(s)$ so that all poles of the **closed loop** transfer function lie in the open left half plane (i.e. the real part of the poles are strictly less than zero, $\{\text{Re}(p_k) < 0 : k = 1, 2, \ldots, n\}$). The **closed loop** transfer function between the reference signal, $r(t)$, and the output, $y(t)$, is

$$T(s) = \frac{G(s)C(s)}{1 + G(s)C(s)} \tag{2.2.1}$$

and the closed loop poles are the values of s that satisfy

$$1 + G(s)C(s) = 0 \tag{2.2.2}$$

so given the transfer function of the system, $G(s)$, the controller transfer function, $C(s)$, has to be designed such that these values of s all lie in the open left half plane.

To determine the stability of the closed loop system in Figure 2.1, we need to check that none of the closed loop poles, i.e. the values of s satisfying

$1 + G(s)C(s) = 0$ lie in the closed right half plane, i.e. values of s where $\mathrm{Re}(s) \geq 0$.

To simplify the notation, define the **open loop** transfer function as

$$L(s) = G(s)C(s) \tag{2.2.3}$$

and write $L(s)$ as the ratio of two polynomials

$$L(s) = \frac{N(s)}{D(s)} \tag{2.2.4}$$

The closed loop poles occur at the values of s that satisfy

$$1 + L(s) = \frac{D(s) + N(s)}{D(s)} = 0 \tag{2.2.5}$$

Hence the closed loop poles occur at the roots of $D(s) + N(s) = 0$. To determine whether any of these poles lie in the right half plane, we use ideas from the residue theorem of complex contour integration, but it is more straightforward to use a graphical argument to explain the approach.

The **closed loop poles** occur at the values of s (i.e. the roots) of

$$D(s) + N(s) = 0 \tag{2.2.6}$$

so in order to determine whether the **closed loop** system is stable, we need to check that none of these poles lies in the **closed** right half plane (i.e. the right half plane including the imaginary axis). This can be checked directly from the Nyquist plot.

The idea is to consider the transfer function as a *mapping* from the complex plane $s = \sigma + \mathrm{j}\omega$ to a different complex plane, $w = u + \mathrm{j}v$ and we look at how **contours** in the s-plane are mapped to **contours** in the w-plane. Consider a general transfer function that has m zeros and n poles

$$Q(s) = \frac{(s - z_1)(s - z_2)\ldots(s - z_m)}{(s - p_1)(s - p_2)\ldots(s - z_n)} \tag{2.2.7}$$

Under this mapping, a specific point in the s-complex plane s_1 is mapped to w_1, a point in the w-plane

$$w_1 = Q(s_1) = \frac{(s_1 - z_1)(s_1 - z_2)\ldots(s_1 - z_m)}{(s_1 - p_1)(s_1 - p_2)\ldots(s_1 - z_n)} \tag{2.2.8}$$

The argument of w_1 is

$$\begin{aligned}\angle w_1 =& \angle(s_1 - z_1) + \angle(s_1 - z_2) + \cdots + \angle(s_1 - z_m) \\ & - \angle(s_1 - p_1) - \angle(s_1 - p_2) - \cdots - \angle(s_1 - p_n) \\ =& \sum_{k=1}^{m} \angle(s_1 - z_k) - \sum_{k=1}^{n} \angle(s_1 - p_k)\end{aligned} \tag{2.2.9}$$

where $\angle(s_1 - z_k)$ is the argument of the line from the zero z_k to the point s_1. The idea is that as we trace s_1 around the contour in s-plane, we monitor $\angle w_1$, the argument of the corresponding points on the mapping of the contour in the w-plane. The key point is that if the contour in the s-plane **encloses** a zero, say z_1, then as we trace round this contour, the argument $\angle(s_1 - z_k)$ increases by 2π. By contrast, if another zero, say z_2 is not enclosed by the contour, then as we trace the contour in the s-plane, the argument $\angle(s_1 - z_2)$ both increases and decreases, so that the net increase for one complete tracing of the contour is zero. Similarly, if a pole p_1 is enclosed by the contour, as we trace around the contour, $\angle(s_1 - p_1)$ will increase by 2π, but because the effect of the poles are subtracted in (2.2.9), the overall contribution to $\angle w_1$ is -2π. If a pole lies outside the contour, then its net contribution to $\angle w_1$ as the contour is traced will be zero.

This means that if a closed contour in the s-plane is traced in a **clockwise** direction, then when this contour is mapped through a general transfer function, $Q(s)$, the number of **clockwise** encirclements of the origin N by the mapped contour is (HLT, page 17)

$$N = Z - P \tag{2.2.10}$$

where Z is the number of zeros of $Q(s)$ enclosed by the contour and P is the number of poles of $Q(s)$ enclosed by the contour. Note that the negative in front of P arises because the enclosed poles cause the origin to be encircled in an anti-clockwise direction. Also, encirclements must be added algebraically; if there is one clockwise encirclement and one anti-clockwise encirclement, they “add up” to zero encirclements.

We can use this result to test the stability of a closed loop system. For a system to be stable when feedback is applied, we need the zeros of $1 + L(s)$

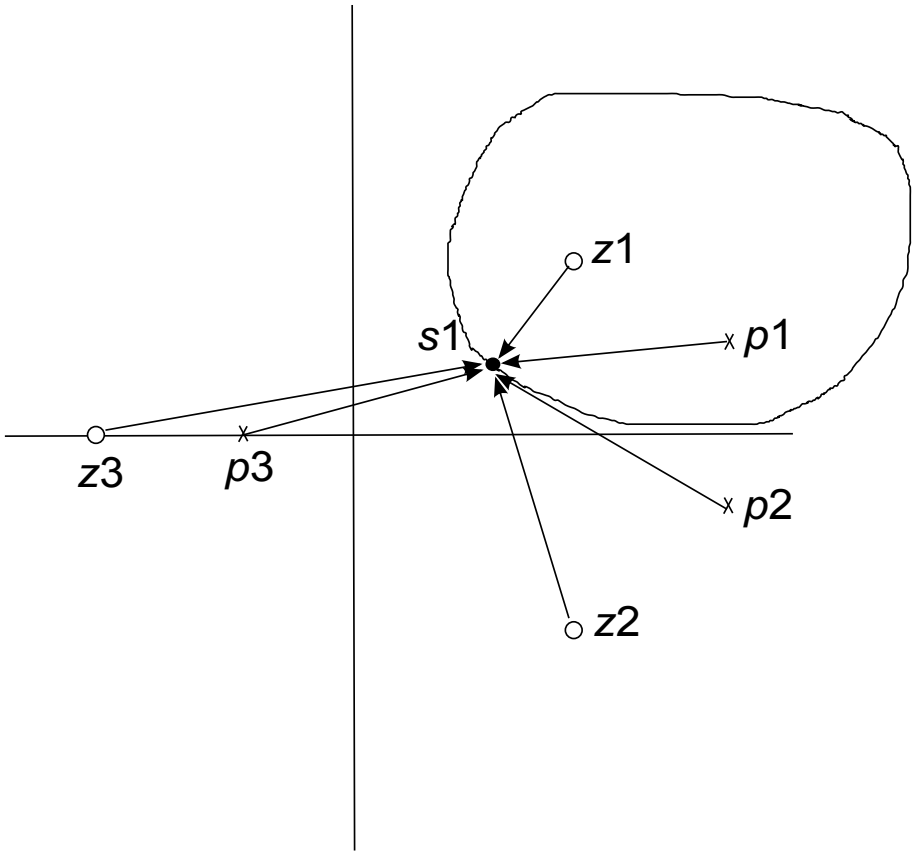

Figure 2.2: Contribution from poles and zeros to mapping of point s_1

to lie in the left half plane. We can turn this around and say that this is equivalent to requiring that none of the zeros of $1 + L(s) = 0$ lie in the right half plane. We can check this by defining a (closed) contour that covers whole of the right half plane, which we map the contour through the transfer function $Q(s) = 1 + L(s)$ and count the number of (clockwise) encirclements of the origin by the mapped contour. For the closed loop system to be stable, we want there to be no zeros enclosed within this mapped contour *i.e.* we require $Z = 0$, so we require that $N = -P$, where P is the number of poles of $1 + L(s)$. However, since

$$1 + L(s) = \frac{D(s) + N(s)}{D(s)} \tag{2.2.11}$$

the *poles* of $1 + L(s)$ are the same as the *poles* of $L(s)$ (which is the **open loop** transfer function without any feedback). This means that P is equal to the number of **open loop** poles lying in the right half plane (*i.e.* the number of **unstable** open loop poles). Often, the system being controlled is open loop stable. For these cases, $P = 0$ and for closed loop stability we require $Z = 0$, so the closed loop system will be stable if the mapped contour does

not encircles the origin, so that $N = 0$.

In order to check stability, we need to map a contour that "encloses" all of the poles and zeros of $1 + L(s)$ that lie in the right half plane. We do this by using the *D*-contour that encloses the "whole" of the right half plane. The *D*-contour consists of three sections (Figure 2.3)

- The positive imaginary axis $s = \mathrm{j}\omega$ from $\omega = 0$ to $\omega = R$, where $R \to \infty$
- The (clockwise) semi-circle $|s| = R$, which has radius R, centred on the origin and goes from $s = R\mathrm{e}^{\mathrm{j}\pi/2}$ to $s = R\mathrm{e}^{-\mathrm{j}\pi/2}$. R is chosen to be large enough to enclose all poles and zeros of $1 + L(s)$ in the right half plane.
- The negative imaginary axis $s = -\mathrm{j}\omega$ from $\omega = R$ to $\omega = 0$

We apply the mapping $1 + L(s)$ to this contour and count the encirclements of the origin.

The final twist is that instead of mapping the *D*-contour though the transfer function $1 + L(s)$ and counting the encirclements of the **origin**, it is easier to map the *D*-contour through $L(s)$ and count the encirclements of the $-1 + \mathrm{j}0$ point. This gives the Nyquist Stability Criterion:

Nyquist Stability Criterion: A closed loop system will be stable if and only if the mapping of the *D*-contour under $L(s)$ encircles the point $-1 + \mathrm{j}0$ in an anti-clockwise direction P times, where P is the number of poles of $L(s)$ in right half plane.

Note. The general form of the Nyquist criterion is given in HLT (p.17), but it is stated in terms of mapping the *D*-contour through a general transfer function $Q(s)$ and counting the encirclements of the *origin*. When checking stability, we count the encirclements of origin when the *D*-contour when mapped through $1 + L(s)$, but it is easier to consider the mapping of the *D*-contour through $L(s)$ and count the encirclements of the $-1 + \mathrm{j}0$ point.

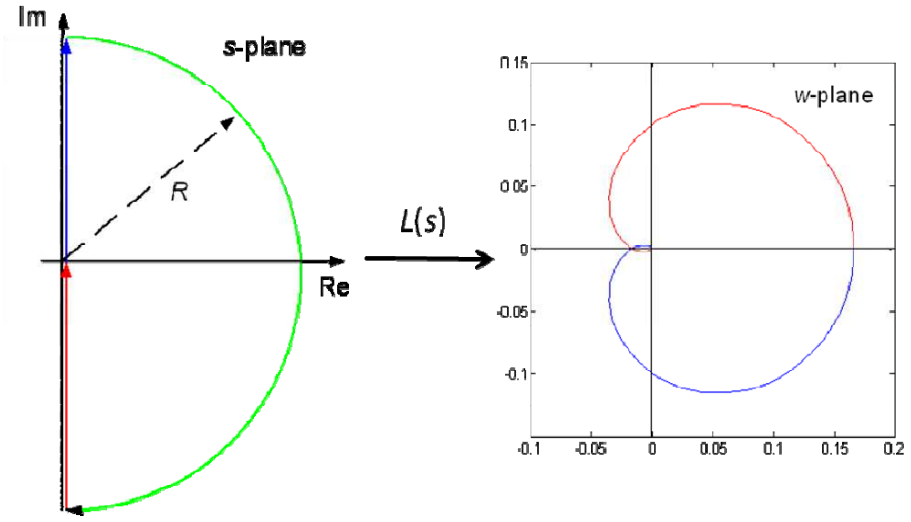

Figure 2.3: *D*-contour and example of the mapping to the *w*-plane

2.3 Drawing Nyquist Plots

The Nyquist plot is the locus of the mapping of the *D*-contour under the transformation $L(s) = G(s)C(s)$. Consider each of the sections of the *D*-contour in turn.

1. The mapping of the **positive** imaginary axis is obtained by evaluating $L(j\omega)$, for $0 \leq \omega \leq R$, which is the **frequency response** of the transfer function $L(s)$ plotted on the complex plane (i.e. on an Argand diagram)

2. The magnitude of the mapping the semi-circle of radius $|s| = R$ for $R \to \infty$ can be written as

$$\lim_{|s|\to\infty} |L(s)| = \lim_{|s|\to\infty} \frac{|N(s)|}{|D(s)|} \tag{2.3.1}$$

   In general, the systems that we will be dealing with are strictly proper which means that the order (*i.e.* the highest power of s in the polynomial) of the numerator polynomial, m, is less than the order of the denominator polynomial, n. This means that

$$\lim_{|s|\to\infty} \frac{|N(s)|}{|D(s)|} = \lim_{|s|\to\infty} \frac{|s|^m}{|s|^n} = 0 \tag{2.3.2}$$

   Hence, in the limit as $R \to \infty$, the semi-circle will map to the **origin** of the complex plane.

3. The mapping of the **negative** imaginary axis from $s = -\mathrm{j}R$ to $s = 0$. This is the complex conjugate of the mapping of the positive real axis, and so can be obtained by reflecting the mapping of $L(\mathrm{j}\omega)$ in the real axis.

This means that drawing a Nyquist plot reduces to plotting the frequency response, $L(\mathrm{j}\omega)$, in the complex plane (*i.e.* on the Argand diagram), since the semi-circle which forms the second segment of the D contour maps to the origin, while the third segment of the Nyquist plot can be obtained by reflecting the frequency response in real axis.

We have assumed that the open loop transfer function $L(s) = G(s)C(s)$ is strictly proper, so that the semi-circle of radius $R \to \infty$ is mapped to the origin of the Nyquist plot. If

$$G(s) = \frac{N_G(s)}{D_G(s)} \qquad C(s) = \frac{N_C(s)}{D_C(s)} \tag{2.3.3}$$

so that

$$L(s) = G(s)C(s) = \frac{N_G(s)N_C(s)}{D_G(s)D_C(s)} \tag{2.3.4}$$

then the order of the numerator polynomial for $L(s)$ is the sum of the orders of the numerator polynomials of $G(S)$ and $C(s)$. Similarly, the order of the denominator polynomial for $L(s)$ is the sum of the orders of the denominator polynomials of $G(S)$ and $C(s)$. This means that $L(s)$ will be strictly proper if one of the transfer functions $G(s)$ or $C(s)$ is strictly proper, even if the other transfer function is only proper.

In practice, it is possible for the controller transfer function $C(s)$ to be proper, rather than strictly proper. For example, the P+I controller

$$C(s) = K_p + \frac{K_i}{s} = \frac{K_p s + K_i}{s} \tag{2.3.5}$$

is proper, but not strictly proper. However, for physical systems, $G(s)$ will always be strictly proper, because physical systems cannot have finite gain at high frequencies. Because the product $G(s)C(s)$ will always be strictly proper, even if $C(s)$ is not, then the closed loop is **well posed** because the open loop transfer function $L(s)$ is strictly proper.

2.4 Example. Unstable Open loop System

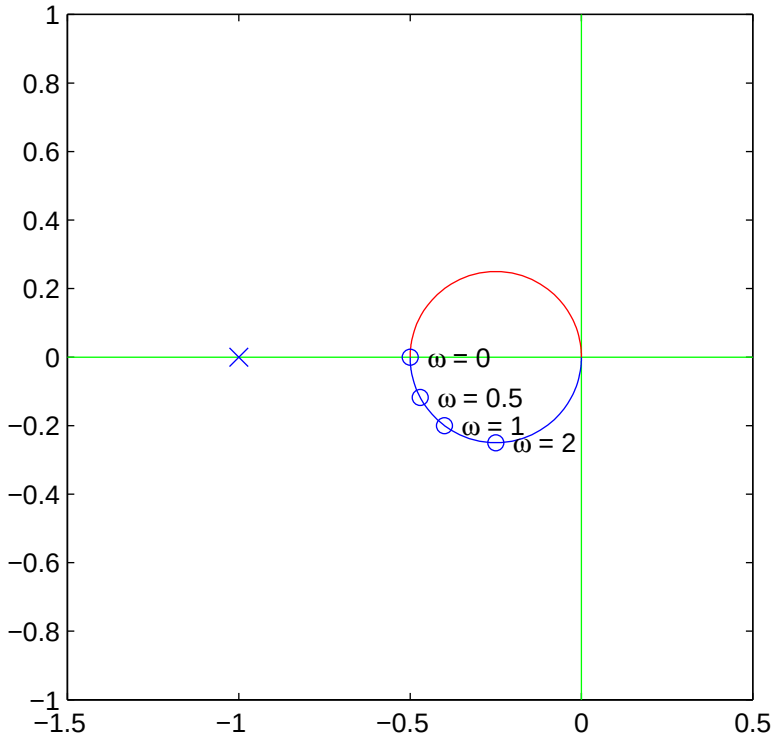

Figure 2.4: Nyquist plot for $L(s) = K/(s-2)$ when $K = 1$

The **open loop** system described by the transfer function

$$L(s) = \frac{K}{s-2} \tag{2.4.1}$$

is unstable because it has a pole in the right half plane at $s = 2$. Figure 2.4 shows the Nyquist plot for the case when $K = 1$. The plot starts at

$$L(0) = -\frac{K}{2} = -0.5 \tag{2.4.2}$$

and because

$$\lim_{\omega\to\infty} L(\mathrm{j}\omega) = \lim_{\omega\to\infty} \frac{K}{\mathrm{j}\omega} = \lim_{\omega\to\infty} -\mathrm{j}\frac{1}{\omega} = 0 \tag{2.4.3}$$

as $\omega \to \infty$, it approaches the origin along the negative imaginary axis. The Nyquist plot does not encircle the $-1 + \mathrm{j}0$ point, but because the **open loop** transfer function has a pole in the right half plane, $P = 1$ and for stability, we require that there is one **anti-clockwise** encirclement of the $-1 + \mathrm{j}0$ point. As a result, the **closed loop** system will be unstable.

If the gain is increased to $K = 2.5$, the Nyquist plot is scaled by a factor of 2.5, as shown in Figure 2.5. This Nyquist plot does encircle the $-1 + j0$ point once in the anti-clockwise direction, so the **closed loop** system will be stable. In this case, applying feedback and increasing the gain K can stabilise an (open loop) unstable system.

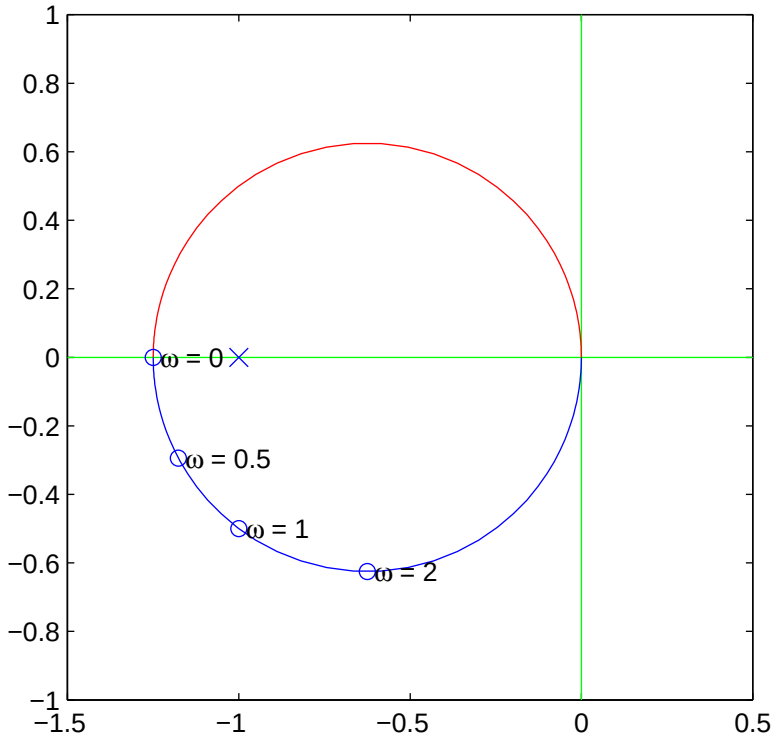

Figure 2.5: Nyquist plot for $L(s) = K/(s-2)$ when $K = 2.5$

2.5 Open Loop Poles and Zeros on Imaginary Axis

We have to be careful when drawing Nyquist plots for systems that have open loop poles and/or zeros that lie on the imaginary axis. For example, consider the open loop system

$$L(s) = \frac{1}{s(s+2)} \tag{2.5.1}$$

which has open loop poles at $s = 0$ and $s = -2$. The Nyquist criterion is based upon counting the encirclements of the origin when the D-contour is

mapped through

$$\begin{aligned} w &= 1 + L(s) \\ &= \frac{s^2 + 2s + 1}{s(s+2)} \\ &= \frac{(s+1)^2}{s(s+2)} \end{aligned} \tag{2.5.2}$$

The argument of the mapped variable is

$$\angle w = 2\angle(s+1) - \angle s - \angle(s+1) \tag{2.5.3}$$

The problem is that the D-contour passes through $s = 0$ and the value of $\angle s$ is undefined at this point.

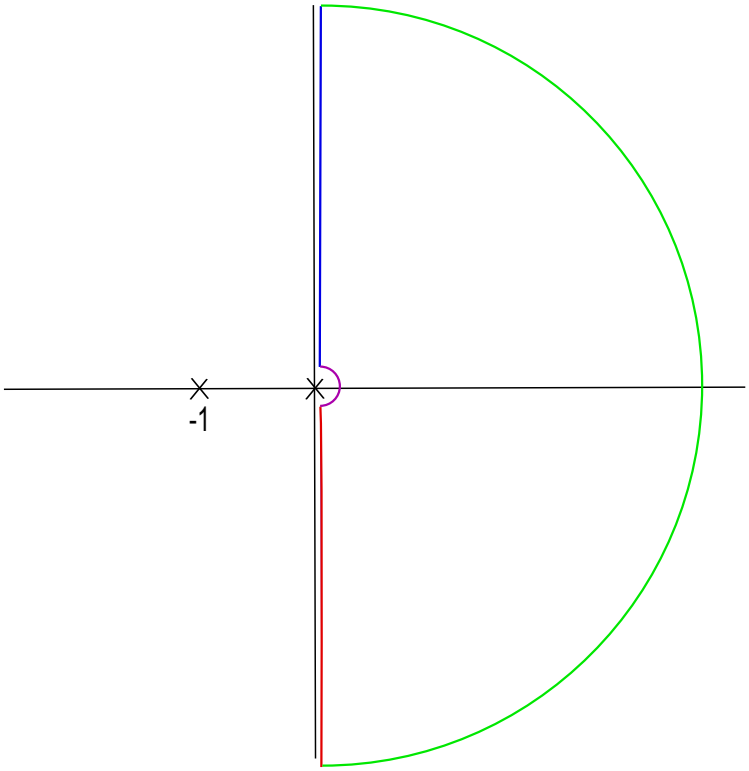

Figure 2.6: D-contour for system with pole at $s = 0$

To overcome this problem, we modify the D-contour so that it includes a small semi-circular indentation of radius ε into the **positive** half of the complex plane, where the semi-circle is centred on the pole. This means that the D-contour has four sections.

1. The mapping of the **positive** imaginary axis is obtained by evaluating $L(\mathrm{j}\omega)$, for $\varepsilon \leq \omega \leq R$, which is the **frequency response** of the transfer

function $L(s)$ plotted on the complex plane, the difference being that in this case, the frequency response starts at $\omega = \varepsilon$ and not at $\omega = 0$.

2. The magnitude of the mapping the semi-circle of radius $|s| = R$ for $R \to \infty$. In the limit as $R \to \infty$, the semi-circle will map to the **origin** of the complex plane.

3. The mapping of the **negative** imaginary axis from $s = -\mathrm{j}R$ to $s = -\mathrm{j}\varepsilon$. This is the complex conjugate of the mapping of the positive real axis, and so can be obtained by reflecting the mapping of $L(\mathrm{j}\omega)$ in the real axis for $-R \leq \omega \leq -\varepsilon$.

4. The mapping of the small semi-circle of radius ε, centred on the origin. This is the mapping of $s = \varepsilon \mathrm{e}^{\mathrm{j}\phi}$ from $\phi = -\pi/2$ to $\phi = \pi/2$, so that

$$w = \frac{\left(\varepsilon \mathrm{e}^{\mathrm{j}\phi} + 1\right)^2}{\varepsilon \mathrm{e}^{\mathrm{j}\phi}\left(\varepsilon \mathrm{e}^{\mathrm{j}\phi} + 2\right)} \tag{2.5.4}$$

For $\varepsilon \to 0$,

$$|w| = \frac{\left|\varepsilon \mathrm{e}^{\mathrm{j}\phi} + 1\right|^2}{\varepsilon \left|\varepsilon \mathrm{e}^{\mathrm{j}\phi} + 2\right|} \tag{2.5.5}$$

$$\to \frac{1}{2\varepsilon} \tag{2.5.6}$$

and

$$\begin{aligned} \angle w &= 2\angle\left(\varepsilon \mathrm{e}^{\mathrm{j}\phi} + 1\right) - \phi - \angle\left(\varepsilon \mathrm{e}^{\mathrm{j}\phi} + 2\right) \\ &\to -\phi \end{aligned} \tag{2.5.7}$$

Hence the mapping of the small semi-circle, is a large semi-circle of radius $1/2\varepsilon$ which starts at $w = \mathrm{j}(1/2\varepsilon)$ and rotates **clockwise** through π radians to $w = -\mathrm{j}(1/2\varepsilon)$. As shown in Fig. 2.7, the inclusion of this semi-circle closes the mapping of the D-contour in the w-plane, allowing us to count the encirclements of the $-1 + \mathrm{j}0$ point.

Although this example has considered a single open loop pole at $s = 0$, the same ideas apply if there are poles at other locations on the imaginary axis, for example at $s = \pm\mathrm{j}\omega$ (note that they will occur in complex conjugate pairs), or if there are zeros on the axis. In each case, the D-contour is modified to include a small semi-circle that takes the contour into the right half plane to avoid the pole or zero.

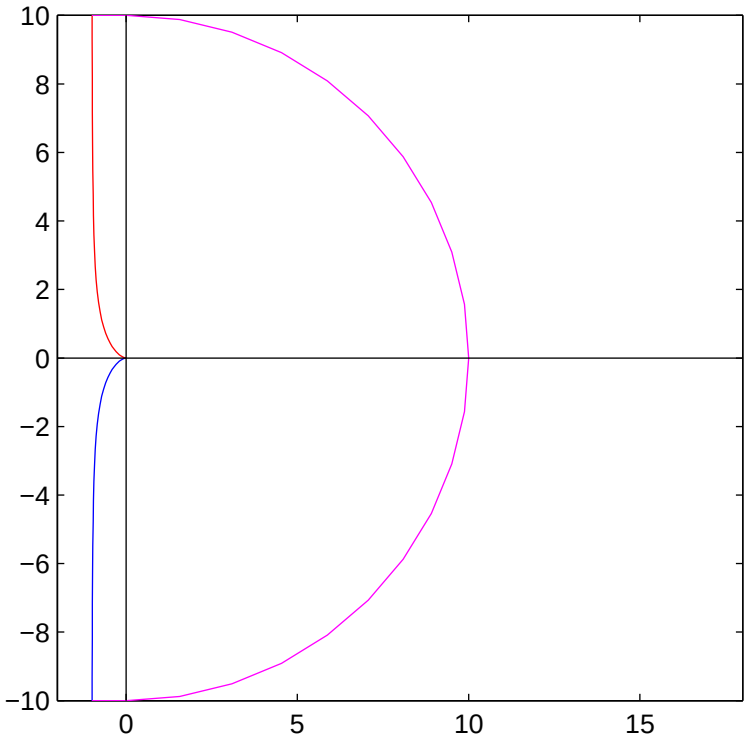

Figure 2.7: Nyquist plot for $L(s) = 1/s(s+1)$

2.6 Interlude: Combining Systems

So far, we have looked at ways of representing single systems, but frequently, we will want to find the transfer function or the state space model of combinations of systems. We will start by considering interconnections of transfer functions and then look at interconnections of state space models

2.6.1 Systems described by Transfer Functions

Suppose that we have two transfer functions

$$G_1(s) = \frac{N_1(s)}{D_1(s)} \qquad G_2(s) = \frac{N_2(s)}{D_2(s)} \tag{2.6.1}$$

then we can obtain expressions for the transfer function, $G_3(s)$, corresponding to the following interconnections (see Figure 2.9).

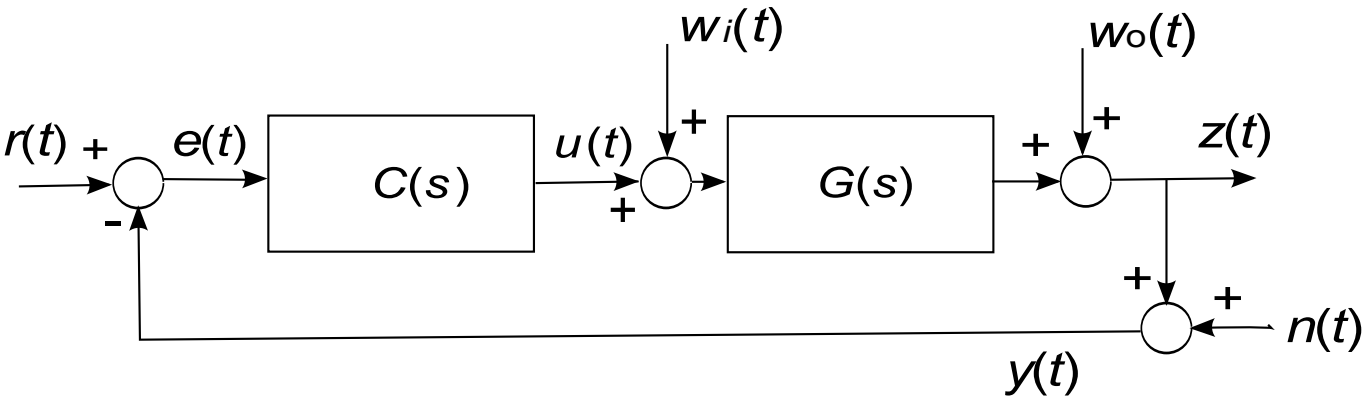

Figure 2.8: Signals in feedback loop

Series If the systems are connected in series then

$$G_3(s) = G_2(s)G_1(s) = \frac{N_2(s)N_1(s)}{D_2(s)D_1(s)} \tag{2.6.2}$$

Parallel Connecting the systems in parallel

$$G_3(s) = G_1(s) + G_2(s) = \frac{N_1(s)}{D_1(s)} + \frac{N_2(s)}{D_2(s)} = \frac{N_1(s)D_2(s) + N_2(s)D_1(s)}{D_2(s)D_1(s)} \tag{2.6.3}$$

Feedback If feedback is applied to $G_1(s)$ with $G_2(s)$ in the feedback loop, then the resulting closed loop transfer function is

$$G_3(s) = \frac{G_1(s)}{1 + G_1(s)G_2(s)} = \frac{N_1(s)D_2(s)}{D_1(s)D_2(s) + N_1(s)N_2(s)} \tag{2.6.4}$$

2.6.2 Systems described by State Space Models

Consider two state space models

$$\begin{aligned} \dot{\mathbf{x}}_1(t) &= \mathbf{A}_1\mathbf{x}_1(t) + \mathbf{b}_1 u_1(t) \\ y_1(t) &= \mathbf{c}_1^{\mathsf{T}}\mathbf{x}_1(t) \end{aligned} \tag{2.6.5}$$

$$\begin{aligned} \dot{\mathbf{x}}_2(t) &= \mathbf{A}_2\mathbf{x}_2(t) + \mathbf{b}_2 u_2(t) \\ y_2(t) &= \mathbf{c}_2^{\mathsf{T}}\mathbf{x}_2(t) \end{aligned} \tag{2.6.6}$$

where for simplicity, we have assumed that both systems are strictly proper, so that $d_1 = d_2 = 0$

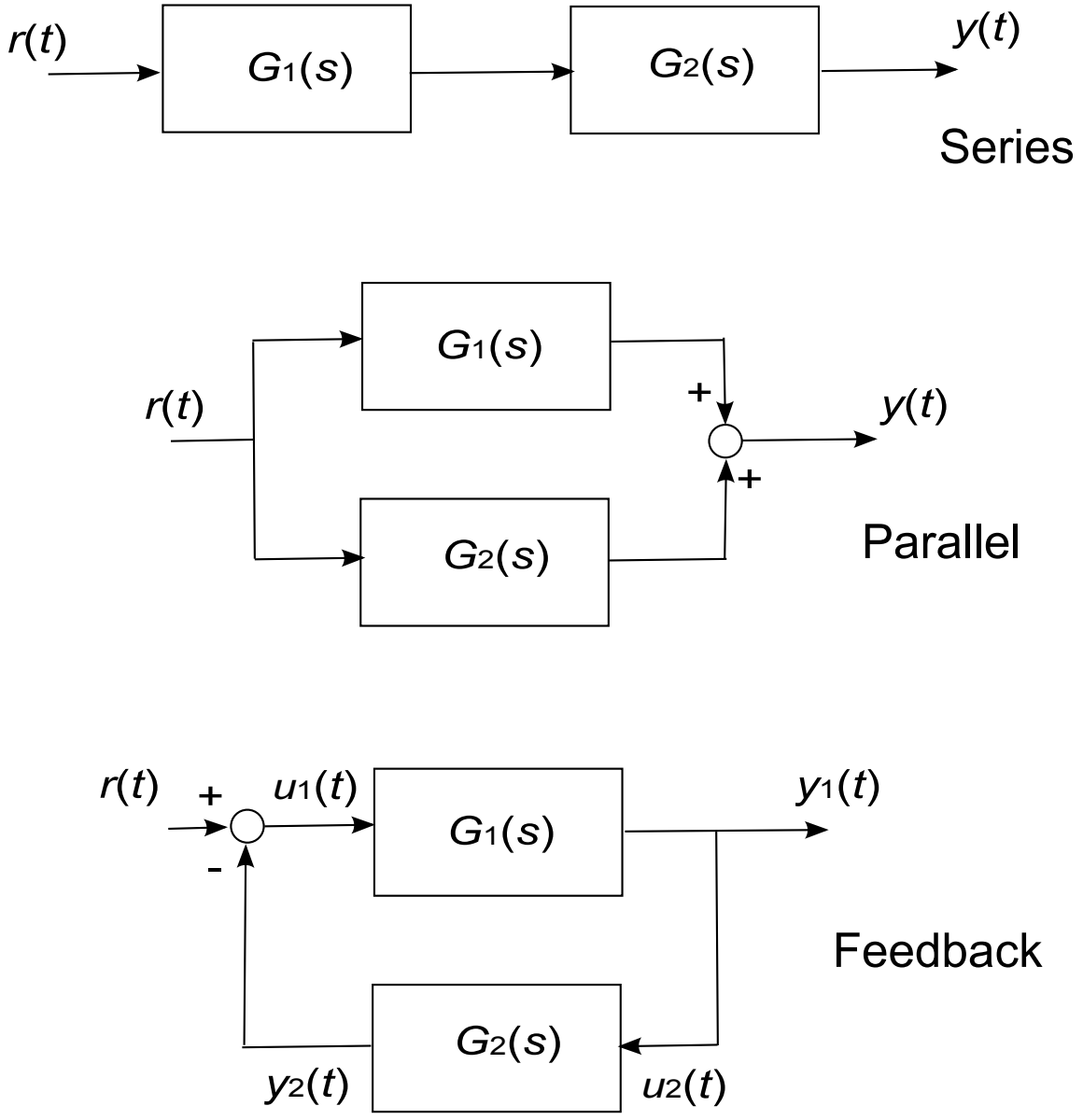

Figure 2.9: Connections of Systems

Series When the two systems are connected in series, the input to the second system is the output from the first system, so that

$$\begin{aligned}\dot{\mathbf{x}}_2(t) &= \mathbf{A}_2\mathbf{x}_2(t) + \mathbf{b}_2 y_1(t) \\ &= \mathbf{A}_2\mathbf{x}_2(t) + \mathbf{b}_2\mathbf{c}_1^{\mathsf{T}}\mathbf{x}_1(t)\end{aligned} \tag{2.6.7}$$

Given that the output of the combined system is the output from the second system, then

$$\begin{bmatrix} \dot{\mathbf{x}}_1(t) \\ \dot{\mathbf{x}}_2(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{b}_2\mathbf{c}_1^{\mathsf{T}} & \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{0} \end{bmatrix} u_1(t) \tag{2.6.8}$$

$$y_2(t) = \begin{bmatrix} \mathbf{0} & \mathbf{c}_2^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} \tag{2.6.9}$$

Defining a combined state vector

$$\mathbf{x}_3(t) = \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} \tag{2.6.10}$$

then

$$\begin{aligned} \dot{\mathbf{x}}_3(t) &= \mathbf{A}_3\mathbf{x}_3(t) + \mathbf{b}_3 u_1(t) \\ y_3(t) &= \mathbf{c}_3^{\mathsf{T}}\mathbf{x}_3(t) \end{aligned} \tag{2.6.11}$$

where

$$\mathbf{A}_3 = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{b}_2\mathbf{c}_1^{\mathsf{T}} & \mathbf{A}_2 \end{bmatrix} \qquad \mathbf{b}_3 = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{0} \end{bmatrix} \qquad \mathbf{c}_3^{\mathsf{T}} = \begin{bmatrix} \mathbf{0} & \mathbf{c}_2^{\mathsf{T}} \end{bmatrix} \tag{2.6.12}$$

Parallel When the two systems are connected in parallel, the output of the combined system is simply the sum of the outputs of the individual systems, so using the combined state vector

$$\mathbf{x}_3(t) = \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} \tag{2.6.13}$$

then

$$\begin{aligned} \dot{\mathbf{x}}_3(t) &= \mathbf{A}_3\mathbf{x}_3(t) + \mathbf{b}_3 u_1(t) \\ y_3(t) &= \mathbf{c}_3^{\mathsf{T}}\mathbf{x}_3(t) \end{aligned} \tag{2.6.14}$$

where

$$\mathbf{A}_3 = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix} \qquad \mathbf{b}_3 = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} \qquad \mathbf{c}_3^{\mathsf{T}} = \begin{bmatrix} \mathbf{c}_1^{\mathsf{T}} & \mathbf{c}_2^{\mathsf{T}} \end{bmatrix} \tag{2.6.15}$$

Feedback If feedback is applied to the first system with $r(t)$ as the reference input, then $u_1(t) = r(t) - y_2(t)$, so that

$$\begin{aligned} \dot{\mathbf{x}}_1(t) &= \mathbf{A}_1\mathbf{x}_1(t) + \mathbf{b}_1\left(r(t) - \mathbf{c}_2^{\mathsf{T}}\mathbf{x}_2(t)\right) \\ &= \mathbf{A}_1\mathbf{x}_1(t) - \mathbf{b}_1\mathbf{c}_1^{\mathsf{T}}\mathbf{x}_2(t) + \mathbf{b}_1 r(t) \end{aligned} \tag{2.6.16}$$

and the output from $G_1(s)$ is

$$y_1(t) = \mathbf{c}_1^{\mathsf{T}}\mathbf{x}_1(t) \tag{2.6.17}$$

Similarly, the output from $G_1(s)$ forms the input to $G_2(s)$, so that

$$\dot{\mathbf{x}}_2(t) = \mathbf{A}_2\mathbf{x}_1(t) + \mathbf{b}_2\mathbf{c}_1^{\mathsf{T}}\mathbf{x}_1(t) \tag{2.6.18}$$

Combining the two systems

$$\begin{bmatrix} \dot{\mathbf{x}}_1(t) \\ \dot{\mathbf{x}}_2(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & -\mathbf{b}_1\mathbf{c}_2^{\mathsf{T}} \\ \mathbf{b}_2\mathbf{c}_1^{\mathsf{T}} & \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{0} \end{bmatrix} r(t) \tag{2.6.19}$$

$$y_1(t) = \begin{bmatrix} \mathbf{c}_1^{\mathsf{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} \tag{2.6.20}$$

and defining a combined state vector

$$\mathbf{x}_3(t) = \begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{bmatrix} \tag{2.6.21}$$

then

$$\begin{aligned} \dot{\mathbf{x}}_3(t) &= \mathbf{A}_3\mathbf{x}_3(t) + \mathbf{b}_3 u_1(t) \\ y_1(t) &= \mathbf{c}_3^{\mathsf{T}}\mathbf{x}_3(t) \end{aligned} \tag{2.6.22}$$

where

$$\mathbf{A}_3 = \begin{bmatrix} \mathbf{A}_1 & -\mathbf{b}_1\mathbf{c}_2^{\mathsf{T}} \\ \mathbf{b}_2\mathbf{c}_1^{\mathsf{T}} & \mathbf{A}_2 \end{bmatrix} \qquad \mathbf{b}_3 = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{0} \end{bmatrix} \qquad \mathbf{c}_3^{\mathsf{T}} = \begin{bmatrix} \mathbf{c}_1^{\mathsf{T}} & \mathbf{0} \end{bmatrix} \tag{2.6.23}$$

2.6.3 MATLAB Commands

In principle, it is possible to work out combinations of systems by hand, but fortunately MATLAB provides the following commands to do this for us.

Series `mm3 = series(mm1,mm2);`

Parallel `mm3 = parallel(mm1,mm2);`

Feedback `mm3 = feedback(mm1,mm2);`

where `mm1` and `mm2` are model structures created by the `tf`, `zpk` or `ss` commands.

Note. If we are only considering unity feedback, then `mm2` is set to 1. It also possible to create the model for a positive feedback connection; type `help feedback` at the MATLAB prompt for more details.

2.7 Internal Stability

For the feedback system in Figure 2.8, there are four "input" signals to the loop, namely $r(t)$, $w_i(t)$, $w_o(t)$ and $n(t)$. There are four "output" signals, $z(t)$, $y(t)$, $e(t)$ and $u(t)$, but because $y(t) = z(t) + n(t)$ and $e(t) = r(t) - y(t) = r(t) - z(t) - n(t)$, only two of these signals, $z(t)$ and $u(t)$ are independent. Since

$$\begin{aligned} Z(s) =& \frac{G(s)C(s)}{1+G(s)C(s)}R(s) + \frac{G(s)}{1+G(s)C(s)}W_i(s) + \frac{1}{1+G(s)C(s)}W_o(s) \\ & - \frac{G(s)C(s)}{1+G(s)C(s)}N(s) \\ U(s) =& \frac{C(s)}{1+G(s)C(s)}R(s) - \frac{G(s)C(s)}{1+G(s)C(s)}W_i(s) - \frac{C(s)}{1+G(s)C(s)}W_o(s) \\ & - \frac{C(s)}{1+G(s)C(s)}N(s) \end{aligned} \tag{2.7.1}$$

which can be written more compactly as

$$\begin{bmatrix} Z(s) \\ U(s) \end{bmatrix} = \frac{1}{1+G(s)C(s)} \begin{bmatrix} G(s)C(s) & G(s) & 1 & -G(s)C(s) \\ C(s) & -G(s)C(s) & -C(s) & C(s) \end{bmatrix} \begin{bmatrix} R(s) \\ W_i(s) \\ W_o(s) \\ N(s) \end{bmatrix} \tag{2.7.2}$$

For this system to be stable, then we require that all of these transfer functions must be stable.

Example 1. Consider a system where

$$G(s) = \frac{s-1}{s+1} \qquad C(s) = \frac{1}{s-1} \tag{2.7.3}$$

so that

$$\begin{aligned} Z(s) &= \frac{G(s)C(s)}{1+G(s)C(s)}R(s) \\ &= \frac{\frac{s-1}{s+1}\frac{1}{s-1}}{1+\frac{s-1}{s+1}\frac{1}{s-1}}R(s) \\ &= \frac{1}{s+2}R(s) \end{aligned} \tag{2.7.4}$$

which means that the transfer function from $R(s)$ to $Z(s)$ is stable. However,

$$\begin{aligned} U(s) &= \frac{C(s)}{1 + G(s)C(s)} R(s) \\ &= \frac{\frac{1}{s-1}}{1 + \frac{s-1}{s+1}\frac{1}{s-1}} R(s) \\ &= \frac{s+1}{(s-1)(s+2)} R(s) \end{aligned} \tag{2.7.5}$$

so the transfer function from $R(s)$ to $U(s)$ is unstable. This means that if a change is applied to the reference signal $r(t)$, the output $z(t)$ will remain bounded, but the input to the system $u(t)$ is unbounded. This is a consequence of using an unstable controller.

Example 2. Now consider a system where

$$G(s) = \frac{1}{s-1} \qquad C(s) = \frac{s-1}{s+1} \tag{2.7.6}$$

In this case, a stable controller is controlling an unstable system and

$$\begin{aligned} Z(s) &= \frac{G(s)C(s)}{1 + G(s)C(s)} R(s) \\ &= \frac{\frac{1}{s-1}\frac{s-1}{s+1}}{1 + \frac{1}{s-1}\frac{s-1}{s+1}} R(s) \\ &= \frac{1}{s+2} R(s) \end{aligned} \tag{2.7.7}$$

so the transfer function from $R(s)$ to $Z(s)$ is stable. Also

$$\begin{aligned} U(s) &= \frac{C(s)}{1 + G(s)C(s)} R(s) \\ &= \frac{\frac{s-1}{s+1}}{1 + \frac{1}{s-1}\frac{s-1}{s+1}} R(s) \\ &= \frac{1}{s+2} R(s) \end{aligned} \tag{2.7.8}$$

so the transfer function from $R(s)$ to $U(s)$ is also stable. However,

$$\begin{aligned} Z(s) &= \frac{G(s)}{1 + G(s)C(s)} W_i(s) \\ &= \frac{\frac{1}{s-1}}{1 + \frac{1}{s-1}\frac{s-1}{s+1}} W_i(s) \\ &= \frac{s+1}{(s-1)(s+2)} W_i(s) \end{aligned} \tag{2.7.9}$$

so the transfer function from $W_i(s)$ to $Z(s)$ is unstable and any disturbance at the input to the system will make the closed loop unstable, making the output $z(t)$ unbounded.

These examples show that it is not sufficient to check the stability from the reference signal $r(t)$ to the output $z(t)$ and in practice we require the more restrictive condition of **internal stability**. For single-input, single-output systems, this can be expressed as

A feedback system is **internally stable** if (and only if) both the following conditions hold

1. The Nyquist stability condition holds so that the transfer function $1 + G(s)C(s)$ has no zeros in the closed right half plane (i.e. $\mathrm{Re}(s) \geq 0$).
2. When the product $G(s)C(s)$ is formed, there are no pole-zero cancelations in the closed right half plane.

Lecture 3

Loop Shaping

Topics

1. Using Bode diagrams
2. Gain margin and phase margin in Bode diagrams
3. Shaping to open loop transfer function to achieve required frequency response for sensitivity and complementary sensitivity
4. Limitations

Learning Outcomes

1. Understand how to interpret the Nyquist stability criterion from a Bode plot
2. Know how to determine the gain margin and phase margin from a Bode plot
3. Understand how performance specifications can be interpreted in terms of sensitivity and complementary sensitivity
4. Understand how shaping the open loop response can achieve a desired specification

5. Understand the relationship between the transfer functions of sensitivity and complementary sensitivity
6. Understand how the Bode integral limits the achievable frequency response of the sensitivity

3.1 Using Bode Diagrams

The Nyquist stability criterion was derived by considering the Nyquist plot, which is the mapping of the D-contour through the open loop transfer function, $L(s)$. Closed loop stability is checked by counting the number of encirclements of the $-1 + \mathrm{j}0$ point by the Nyquist plot. The Nyquist plot effectively is a plot of the frequency response of the open loop transfer function, $L(\mathrm{j}\omega)$, on the complex plane (i.e. the Argand diagram). Although the Nyquist plot is primarily used for checking stability, it can be used as the basis of a controller design that achieves a given specification on performance such as closed loop bandwidth, particularly because we can use M-circles to determine the closed loop frequency response directly from the Nyquist plot. However, given that the Nyquist diagram is a plot of the frequency response, there are advantages to using a Bode diagram to represent $L(\mathrm{j}\omega)$.

The Bode diagram plots $20\log_{10}|L(\mathrm{j}\omega)|$ the magnitude of the frequency response in dB, and $\angle L(\mathrm{j}\omega)$, the argument of $L(\mathrm{j}\omega)$ in degrees against $\log_{10}\omega$. The main benefit of using the Bode diagram is that because $L(s) = G(s)C(s)$, then

$$\begin{aligned} 20\log_{10}|L(\mathrm{j}\omega)| &= 20\log_{10}|G(\mathrm{j}\omega)| + 20\log_{10}|C(\mathrm{j}\omega)| \\ \angle L(\mathrm{j}\omega) &= \angle G(\mathrm{j}\omega) + \angle C(\mathrm{j}\omega) \end{aligned} \tag{3.1.1}$$

This means that the effect of including a compensator (controller) with transfer function $C(s)$ by summing the magnitude and phase Bode plots.

MATLAB will provide Bode plots for you. For a system `m`, which can be in `tf`, `zpk` or `ss` format,

- `bode(m);` plots the Bode plot of the system

- the `bode` command selects an appropriate frequency range for the plot, but sometimes, particularly if the system contains integrators, the frequency range can be so large that range on the magnitude and phase plots are also large, making it difficult to see the details. The commands

```
w = logspace(-2,3,50);
bode(m,w);
grid on
```

  generates a vector 50 frequencies that are equally spaced on a logarithmic scale between $\omega = 10^{-2}$ and $\omega = 10^{3}$ and then creates the Bode plot at these frequencies. Adjusting the arguments in the `logspace` command allows you to choose the frequency range for the Bode plot and the number of frequencies used to produce the plot. The command `grid on` draws a grid on the plot.

- `[mag,phase,w] = bode(m);` returns the magnitude and phase of the response at the frequencies in the vector `w`, but does not produce the Bode plot. Note that `mag` is **not** in dB. Also, because `bode` is set up to handle multi-input, multi-output systems, `mag` and `phase` are returned as 3-dimensional arrays, although for a SISO system, the size of the first two dimensions is unity.

A reminder of some important features of Bode plots

- If $L(s)$ is expressed in pole-zero form, then a reasonable approximation to the Bode plot can be obtained by "adding-up" the asymptotic approximations for magnitude and phase plots of each of the zeros and poles. For complex conjugate poles, we need to use the frequency response of the standard 2nd order system as given in HLT p.167.

- Because we have assumed that the closed loop is **well posed**, then $L(s)$ is strictly proper, so the magnitude plot rolls-off at high frequencies at a rate $-20(n-m)$ dB/decade, where m and n are the orders of the numerator and denominator polynomial respectively.

- If there is no time delay in the system, then at high frequencies ($\omega \to \infty$), $\angle L(\mathrm{j}\omega) \to -90^{\mathrm{o}} \times (n-m)$

- If there are no integrators, then at low frequencies ($\omega \to 0$), $20\log_{10}|L(\mathrm{j}\omega)|$ approaches the DC gain (measured in dB) and $\angle L(\mathrm{j}\omega) \to 0^{\mathrm{o}}$.
- If the open loop includes one (or more) integrators, then low frequency ($\omega \to 0$) asymptote of the magnitude plot is $-20k$ dB/decade, where k is the number of integrators. At low frequencies, the phase plot approaches $-90^{\mathrm{o}} \times k$.
- If there is a time delay τ, so that

$$L(s) = L_1(s)\mathrm{e}^{-s\tau} \tag{3.1.2}$$

then

$$|L(\mathrm{j}\omega)| = |L_1(\mathrm{j}\omega)| \qquad \angle L(\mathrm{j}\omega) = \angle L_1(\mathrm{j}\omega) - \omega\tau \tag{3.1.3}$$

This means that adding a time delay to the system (or to the controller) does not affect the magnitude plot, but the phase lag of $\angle L(\mathrm{j}\omega)$ is increased (i.e. made more negative) by $\omega\tau$. Note that although the phase lag increases linearly with increasing ω, the shift will not be linear on the phase plot as $\angle L(\mathrm{j}\omega)$ is plotted against $\log_{10}\omega$

Example 1. The Bode plot of

$$L(s) = \frac{100(s+5)}{(s+1)(s+10)(s+30)} \tag{3.1.4}$$

is shown in Figure 3.1. Because there are 1 zero and 3 poles, at high frequencies the magnitude plot rolls off at -40dB per decade and the phase plot approaches -180^{o}. The DC (low frequency gain of the system is $20\log_{10}(500/300) = 4.44$dB.

Example 2. If a delay of 0.005s is included in the system, then the transfer function becomes

$$L(s) = \frac{100(s+5)}{(s+1)(s+10)(s+30)}\mathrm{e}^{-0.005s} \tag{3.1.5}$$

is shown in Figure 3.2. Adding the time delay does not alter the magnitude plot, but the phase becomes more negative at high frequencies.

Example 3. The Bode plot of

$$L(s) = \frac{100(s+5)}{s(s+1)(s+10)(s+30)} \tag{3.1.6}$$

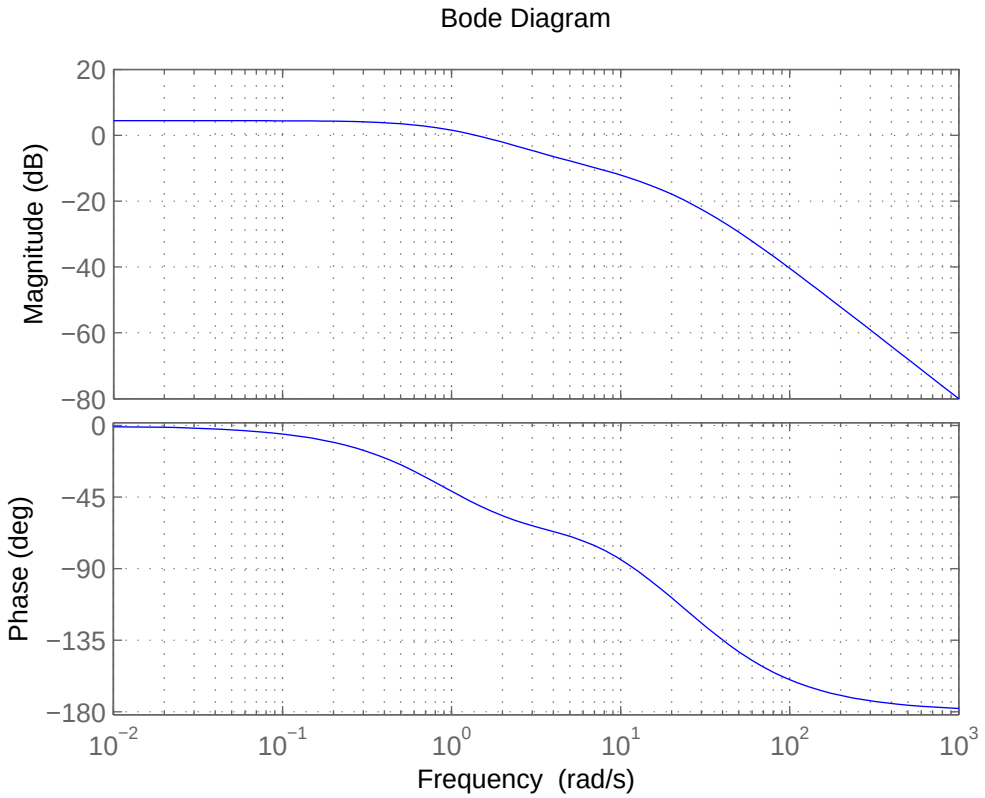

Figure 3.1: Bode plot of $\frac{100(s+5)}{(s+1)(s+10)(s+30)}$

where an integrator has been added to the transfer function in Example 1 is shown in Figure 3.3. In this case, there is still one zero, but the number of poles has increased to 4, so the magnitude plot rolls off at -60dBb per decade at high frequencies, while the phase plot approaches -270^{o}. Because the transfer function has an integrator, at low frequencies the gain of the system approaches ∞ and the slope of magnitude plot is -20dBb per decade, while the phase plot starts at -90^{o}.

3.1.1 Checking Closed Loop Stability from Bode Plot

If the open loop system is stable, then the closed loop will be stable if the Nyquist plot does not encircle the $-1 + \text{j}0$ point. Usually, this means that if the Nyquist plot crosses the imaginary axis at ω_{G}, so that $\angle L(\text{j}\omega_{\text{G}}) = -180^{\text{o}}$, then for stability, we require $|L(\text{j}\omega_{\text{G}})| < 1$. On the Bode plot, this corresponds to requiring that

$$20 \log_{10} |L(\text{j}\omega_{\text{G}})| < 0\text{dB} \tag{3.1.7}$$

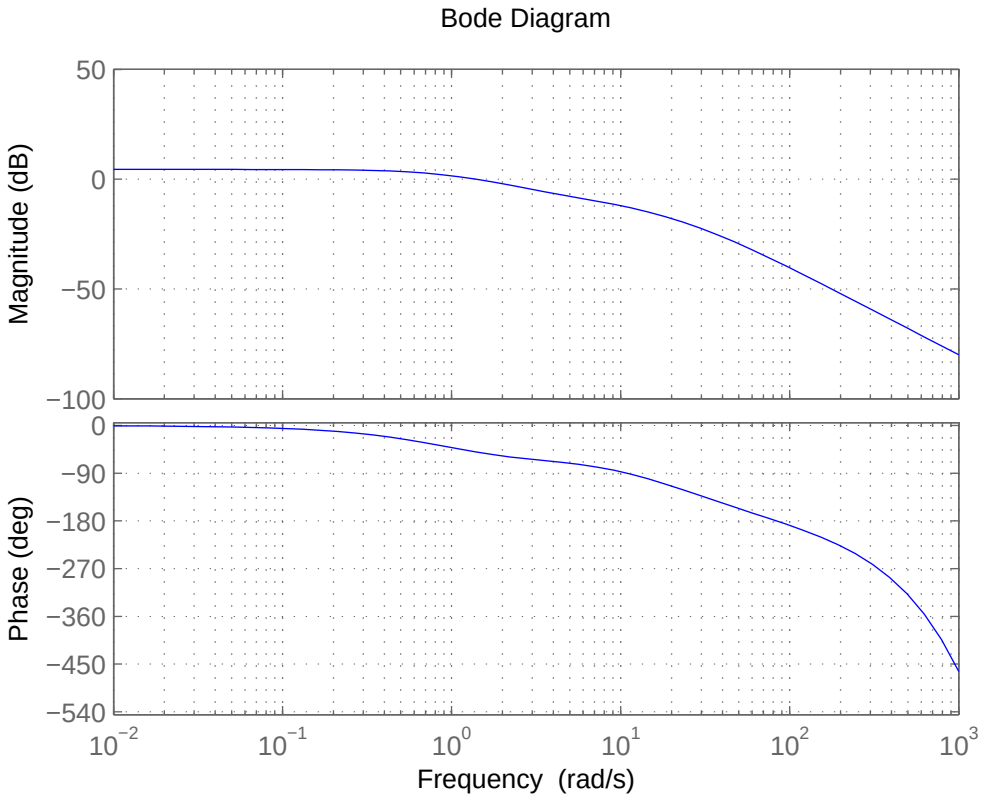

Figure 3.2: Bode plot of $\frac{100(s+5)}{(s+1)(s+10)(s+30)}e^{-0.005s}$

which means that the magnitude plot must be below 0dB before (i.e. at a lower frequency) the phase plot crosses -180^o.

Note Determining stability from the Bode plot for cases where the open loop is unstable so that it is necessary to count encirclements of the $-1 + j0$ point is possible, but is tricky (the idea is that you track the magnitude as the phase plot crosses -180^o). In these cases, it is usually better to use the Nyquist plot to determine stability. As we will see below, the Bode plot is better suited as a tool for designing a controller that achieves a given level of performance, rather than for assessing stability.

3.1.2 Gain and Phase Margins

We can also use the Bode plot to determine the gain margin and phase margin of the closed loop system. Recall that the gain margin is the amount by which the gain of the open loop system can be increased before the closed loop system becomes unstable. If the Nyquist plot crosses the negative real

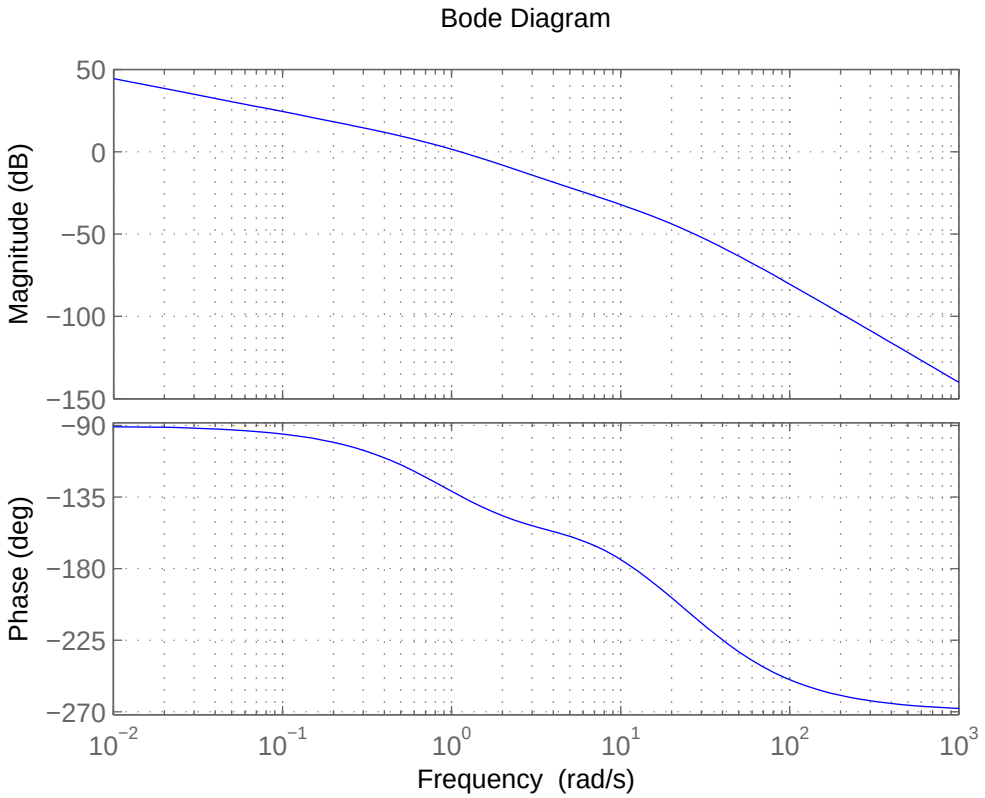

Figure 3.3: Bode plot of $\frac{100(s+5)}{s(s+1)(s+10)(s+30)}$

axis at frequency ω_G, so that $\angle L(j\omega_G) = -180^o$, then the gain margin is

$$GM = \frac{1}{|L(j\omega_G)|} \tag{3.1.8}$$

If we express the GM in dB, then

$$20\log_{10} GM = -20\log_{10} |L(j\omega_G)| \tag{3.1.9}$$

Note that if the closed loop system is stable, then $|L(j\omega_G)| < 1$, so $20\log_{10}|L(j\omega_G)| < 0$, which means that $20\log_{10} GM$ will be positive. Using the Bode plot, we can identify ω_G as the frequency where the phase plot crosses -180^o and we can then read $20\log_{10}|L(j\omega_G)|$ directly from the magnitude plot, as shown in Figure 3.4 for the system in Example 3.

The phase margin is the amount of phase lag that can be introduced into the open loop before the closed loop becomes unstable. If ω_P is the frequency at which the open loop gain is unity, so that $|L(j\omega_P)| = 1$, then the phase margin is the angle between $\angle L(j\omega_P)$ and the negative real axis, which means that

$$PM = 180^o + \angle L(j\omega_P) \tag{3.1.10}$$

Again this can be read directly from the Bode plot, as shown in Figure 3.4. Identify ω_P as the frequency that the magnitude plot crosses 0dB and then from the phase plot, read off the difference in angle between $\angle L(j\omega_P)$ and -180^o.

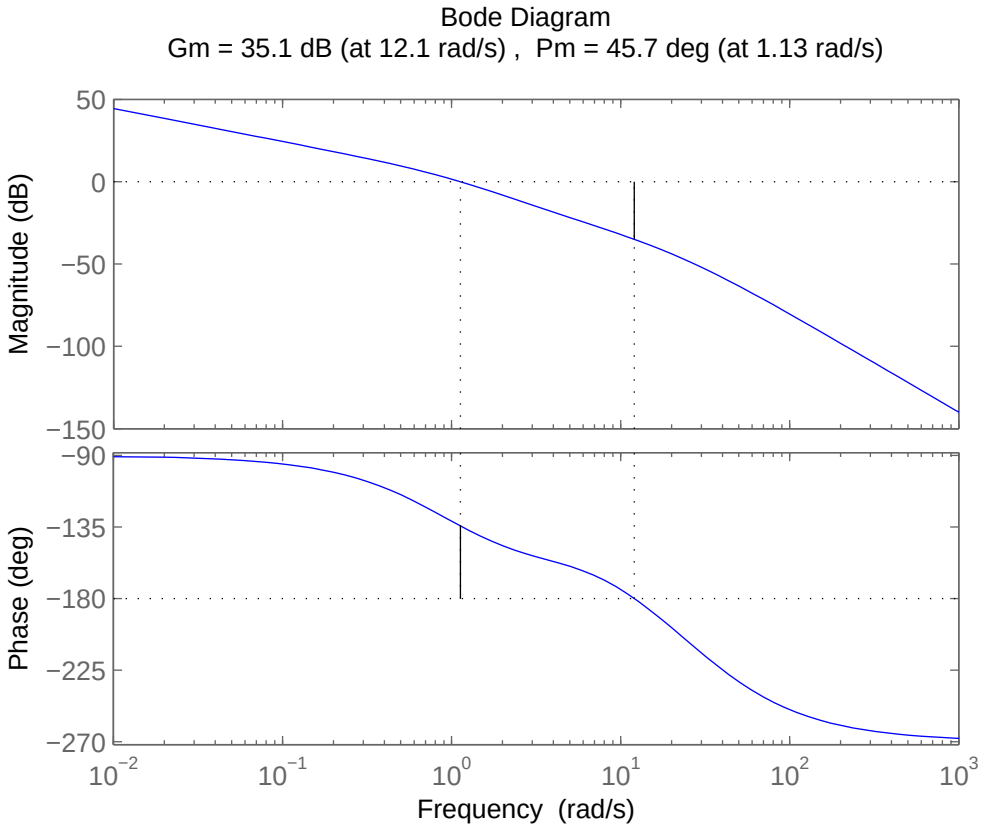

Figure 3.4: Bode plot of $\frac{100(s+5)}{s(s+1)(s+10)(s+30)}$ showing gain and phase margins

MATLAB provides a function that gives the gain and phase margins for a system. For a system `m`

- `margin(m);` produces the Bode plot which includes the gain and phase margins together with the frequencies ω_G and ω_P
- `[Gm,Pm,Wg,Wp] = margin(m);` returns the gain and phase margins and ω_G and ω_P, but does not produce the Bode plot.

3.2 Closed Loop Performance

For the system in Figure 3.5, the closed loop transfer functions from the reference signal $r(t)$, the disturbance $d(t)$ and the sensor noise $n(t)$ to the

output $z(t)$ are

$$Z(s) = \frac{G(s)C(s)}{1 + G(s)C(s)}R(s) + \frac{1}{1 + G(s)C(s)})D(s) - \frac{G(s)C(s)}{1 + G(s)C(s)}N(s) \quad (3.2.1)$$

Define the **sensitivity transfer function**

$$S(s) = \frac{1}{1 + G(s)C(s)} \quad (3.2.2)$$

and the **complementary sensitivity transfer function**

$$T(s) = \frac{G(s)C(s)}{1 + G(s)C(s)} \quad (3.2.3)$$

so that

$$Z(s) = T(s)R(s) + S(s)D(s) - T(s)N(s) \quad (3.2.4)$$

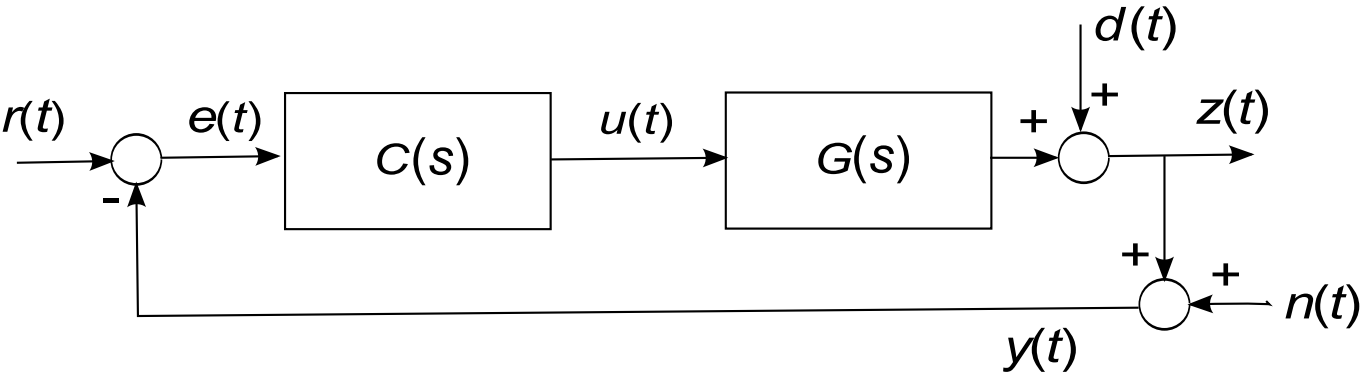

Figure 3.5: Signals in closed loop system

This means that the frequency responses are

$$\begin{aligned} Z(\mathrm{j}\omega) &= \frac{G(\mathrm{j}\omega)C(\mathrm{j}\omega)}{1 + G(\mathrm{j}\omega)C(\mathrm{j}\omega)}R(\mathrm{j}\omega) + \frac{1}{1 + G(\mathrm{j}\omega)C(\mathrm{j}\omega)}D(\mathrm{j}\omega) - \frac{G(\mathrm{j}\omega)C(\mathrm{j}\omega)}{1 + G(\mathrm{j}\omega)C(\mathrm{j}\omega)}N(\mathrm{j}\omega) \\ &= T(\mathrm{j}\omega)R(\mathrm{j}\omega) + S(\mathrm{j}\omega)D(\mathrm{j}\omega) - T(\mathrm{j}\omega)N(\mathrm{j}\omega) \end{aligned} \quad (3.2.5)$$

We want the controlled system to “reject” the effects of the disturbance, so that the output $z(t)$ is not affected by the disturbance $d(t)$. Since

$$|Z(\mathrm{j}\omega)| = |S(\mathrm{j}\omega)|\,|D(\mathrm{j}\omega)| \quad (3.2.6)$$

this means that we want $|S(j\omega)|$ to be small. We also want the closed loop to "track" changes in the reference signal and since

$$|Z(j\omega)| = |T(j\omega)|\,|R(j\omega)| \tag{3.2.7}$$

we need $|T(j\omega)|$ to be unity.

There are a number a number of reasons why we cannot completely achieve the goal of making $|S(j\omega)|$ small and $|T(j\omega)| = 1$.

1. Because we have assumed that the closed loop system is **well-posed**, the open loop transfer function $L(s) = G(s)C(s)$ is strictly proper, which means that

$$\begin{aligned}\lim_{\omega\to\infty} |L(j\omega)| &= \lim_{\omega\to\infty} |G(j\omega)C(j\omega)| \\ &= 0\end{aligned} \tag{3.2.8}$$

   This means that

$$\begin{aligned}\lim_{\omega\to\infty} |T(j\omega)| &= \lim_{\omega\to\infty} \frac{|G(j\omega)C(j\omega)|}{|1 + G(j\omega)C(j\omega)|} \\ &= 0\end{aligned} \tag{3.2.9}$$

   so $z(t)$ cannot track $r(t)$ at high frequencies.

2. From (3.2.2) and (3.2.3)

$$\begin{aligned}S(s) + T(s) &= \frac{1}{1 + G(s)C(s)} + \frac{G(s)C(s)}{1 + G(s)C(s)} \\ &= 1\end{aligned} \tag{3.2.10}$$

   Using (3.2.9)

$$\begin{aligned}\lim_{\omega\to\infty} |S(j\omega)| &= \lim_{\omega\to\infty} |1 - T(j\omega)| \\ &\leq \lim_{\omega\to\infty} \{1 + |T(j\omega)|\} \\ &= 1\end{aligned} \tag{3.2.11}$$

   so we cannot make $|S(j\omega)|$ small at high frequencies, which means that the disturbance will affect the output at these frequencies.

3. From (3.2.5), the magnitude of the frequency response between the sensor noise and the output is the same the magnitude of the frequency response from the reference signal to the output. This means that for frequencies where $|T(j\omega)| \approx 1$, so that the output is tracking the reference signal, the effect of the sensor noise at the output is not being attenuated.

These constraints mean that it is not possible to design a control system so that at all frequencies

$$|T(j\omega)| = 1 \text{ so that } z(t) \text{ follows } r(t)$$

and

$$|S(j\omega)| = 0 \text{ so that } d(t) \text{ has no effect on } z(t)$$

and

$$|T(j\omega)| = 0 \text{ so that } n(t) \text{ has no effect on } z(t)$$

As a result, the design of the controller has to be a trade-off between these conflicting requirements. In order to make these trade-offs, we need to know something about the spectra of the signals affecting the system, i.e., $R(j\omega)$, $D(j\omega)$ and $N(j\omega)$. In practice, the spectra of $R(j\omega)$ and $D(j\omega)$ tend to be concentrated at low frequencies, while $n(t)$ is a noise signal, so its spectrum $N(j\omega)$ tends to be concentrated at high frequencies. Choose a frequency ω_c above the bandwidth of both $r(t)$ and $d(t)$, so that

$$|R(j\omega)| < \frac{1}{\sqrt{2}} \quad \text{for } \omega > \omega_c \tag{3.2.12}$$

and

$$|D(j\omega)| < \frac{1}{\sqrt{2}} \quad \text{for } \omega > \omega_c \tag{3.2.13}$$

We then require that the sensor noise is negligible below this frequency, so that

$$|N(j\omega)| \approx 0 \quad \text{for } \omega \leq \omega_c \tag{3.2.14}$$

If this requirement cannot be satisfied, then we have the following options

- Replace the sensor with a better one whose noise characteristics can achieve this constraint.
- Accept that the sensor noise will affect the output
- Reduce ω_c to a frequency where the sensor noise satisfies the constraint in (3.2.14), which means accepting that the tracking of the reference signal by the output will be degraded and/or the disturbance will have an effect on the output.

Note. We have discussed the choice of ω_c in terms of the frequency spectrum of the reference signal $r(t)$, but the requirement for the closed loop system to track the reference signal is usually expressed in terms of the speed of the step response. However, there is a direct relationship between the closed loop bandwidth ω_c and the speed of the step response. The top plot in Figure 3.6 shows the magnitude Bode plot of two closed loop transfer functions, while the lower plot shows the corresponding closed loop step responses. The step response for the closed loop system with the higher bandwidth (dashed) line has the faster response.

Once a suitable value of ω_c has been chosen, then we want to choose $C(s)$ so that $|T(j\omega)| \approx 1$ for $\omega < \omega_c$ so that $z(t)$ tracks $r(t)$, and $|T(j\omega)| \approx 0$ for $\omega > \omega_c$ to ensure that $n(t)$ has no effect on $z(t)$. Since

$$|T(j\omega)| = \frac{|G(j\omega)C(j\omega)|}{|1 + G(j\omega)C(j\omega)|} \tag{3.2.15}$$

these conditions can be satisfied by choosing $C(s)$ so that $|G(j\omega)C(j\omega)|$ is large for $\omega < \omega_c$ and $|G(j\omega)C(j\omega)|$ is small for $\omega > \omega_c$. Referring to Figure 3.7, more formally, we can express this as requiring that

$$|G(j\omega)C(j\omega)| > M_L \quad \text{for } \omega < \omega_L \tag{3.2.16}$$

$$|G(j\omega)C(j\omega)| < \frac{1}{M_H} \quad \text{for } \omega > \omega_H \tag{3.2.17}$$

where $M_L \gg 1$ and $M_H \gg 1$ are constants and $\omega_L < \omega_c$ and $\omega_H > \omega_c$. In addition, if we choose $C(s)$ so that $|G(j\omega_c)C(j\omega_c)| \approx 1$, which means that the Bode plot of $L(s) = G(s)C(s)$ crosses 0 dB at (or close to) ω_c, then

$$|T(j\omega_c)| = \frac{|G(j\omega_c)C(j\omega_c)|}{|1 + G(j\omega_c)C(j\omega_c)|} \approx \frac{1}{2} \tag{3.2.18}$$

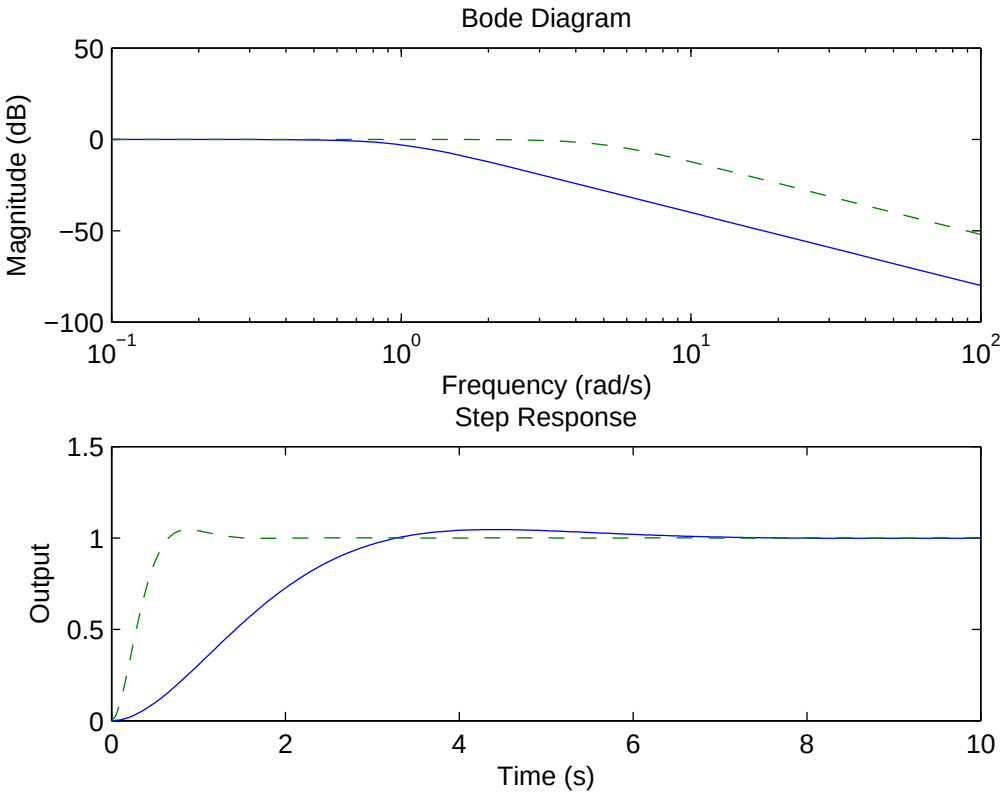

Figure 3.6: Magnitude Bode plots of two closed loop transfer functions (top) with their closed loop step responses (bottom)

which means that ω_c is approximately equal to the closed loop bandwidth (it would be at the closed loop bandwidth if $|T(j\omega_c)| = 1/\sqrt{2}$).

Because ω_c is the frequency at which $G(j\omega)C(j\omega)$ crosses 0 dB, it is referred to as the **crossover** frequency.

If we have designed $C(s)$ so that the low frequency condition in (3.2.16) is satisfied, then

$$|S(j\omega)| = \frac{1}{|1 + G(j\omega)C(j\omega)|} \approx 0 \quad \text{for } \omega < \omega_L \tag{3.2.19}$$

which ensures that the effect of the disturbances on the output is small over the frequency range. However,

$$|S(j\omega)| = \frac{1}{|1 + G(j\omega)C(j\omega)|} \approx 1 \quad \text{for } \omega > \omega_H \tag{3.2.20}$$

This is a consequence of the constraint that $S(j\omega) + T(j\omega) = 1$, so $|S(j\omega)|$ and $|T(j\omega)|$ can not both be small. This means that is there are any components of the disturbance at frequencies above ω_c, then they will affect the output.

The design process can be thought of as choosing $C(s)$ in order to "shape" the open loop frequency response so that $|G(\mathrm{j}\omega)C(\mathrm{j}\omega)|$ satisfies a given set of constraints. for this reason, the design approach is referred to as **loop shaping**.

Example 4. Suppose that the transfer function of the system is

$$G(s) = \frac{100(s+5)}{s(s+1)(s+10)(s+30)} \tag{3.2.21}$$

The Bode plot for $G(s)$ is shown in Figure 3.7 and it can be seen that the crossover frequency is around 1 rad.s^{-1}. We are given the requirement that

$$20\log_{10}|G(\mathrm{j}\omega)C(\mathrm{j}\omega)| > 30 \text{ dB} \quad \text{for } \omega < 0.1 \text{ rad.s}^{-1} \tag{3.2.22}$$

and

$$20\log_{10}|G(\mathrm{j}\omega)C(\mathrm{j}\omega)| < -40 \text{ dB} \quad \text{for } \omega > 10 \text{ rad.s}^{-1} \tag{3.2.23}$$

These constraints are marked by the red dashed lines in the plot and it can be seen that without any controller (i.e. $C(s) = 1$), the requirements are not met in the regions around 0.1 rad.s^{-1} and around 10 rad.s^{-1}.

If we use a controller with transfer function

$$C(s) = \frac{0.07s + 7}{s + 5} \tag{3.2.24}$$

then from the Bode plot of the controller in Figure 3.8, this is a lag controller with DC gain $7/5 = 2.9$ dB and an attenuation of -23 dB at high frequencies. Including this controller in the feedback loop raises the open loop gain at low frequencies and lowers the gain at high frequencies and as shown in Figure 3.9, the magnitude plot now lies outside the constraint region, so $|G(\mathrm{j}\omega)C(\mathrm{j}\omega)|$ (nearly) satisfies the constraints. However, including the lag controller introduces phase lag, which has the effect of reducing the gain margin from 35.1 dB for the uncompensated system to 11.5 dB with the controller.

The Bode plot of $T(\mathrm{j}\omega)$ is shown in Figure 3.10 and it can be seen that the closed loop bandwidth is around 2 rad.s^{-1}, which is close to ω_c. At frequencies below the closed loop bandwidth, $|T(\mathrm{j}\omega)| = 1$ as required, although there is a resonant peak in $|T(\mathrm{j}\omega)|$, so we would expect there to be some overshoot in the closed loop step response. Above the closed loop bandwidth, $|T(\mathrm{j}\omega)|$ approaches zero. From Figure 3.11, $|S(\mathrm{j}\omega)|$ is low at low frequencies and equals unity at high frequencies.

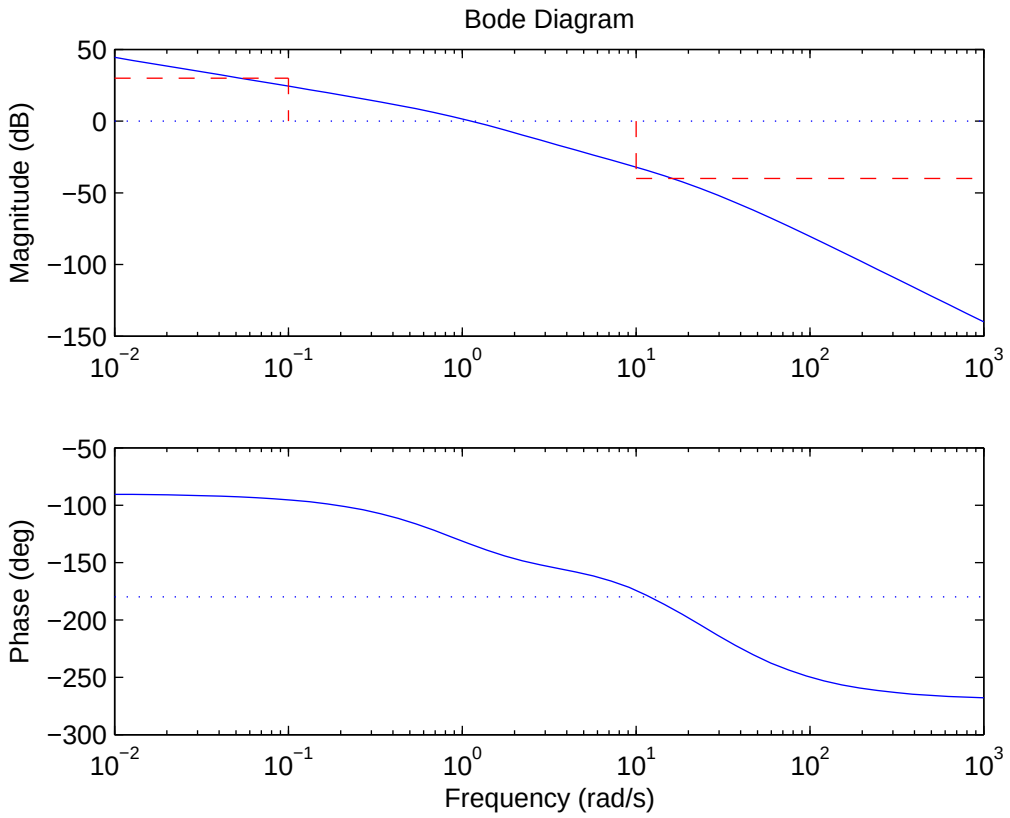

Figure 3.7: Bode plot of $\frac{100(s+5)}{s(s+1)(s+10)(s+30)}$ including design constraints

3.3 Weighting Transfer Functions

For a given ω_c, the values of M_L, M_H, ω_L and ω_H determine the frequency range for the transition from $|G(j\omega)C(j\omega)|$ being large at low frequencies to $|G(j\omega)C(j\omega)|$ being small at high frequencies. These values are design parameters that need to be chosen by the designer of the control system. If $\omega_H - \omega_L$ is small, so that the range of frequencies over which the transition occurs is small, then the gradient of $|G(j\omega)C(j\omega)|$ is high over this region. As a consequence, the change in the phase $\angle G(j\omega)C(j\omega)$ will be large over this region, which reduces the phase margin, making the system susceptible to uncertainties.

An alternative approach is think of the design process as choosing $C(s)$ so that $|S(j\omega)| \ll 1$ for $\omega < \omega_L$ and $|T(j\omega)| \ll 1$ for $\omega > \omega_H$. These conditions

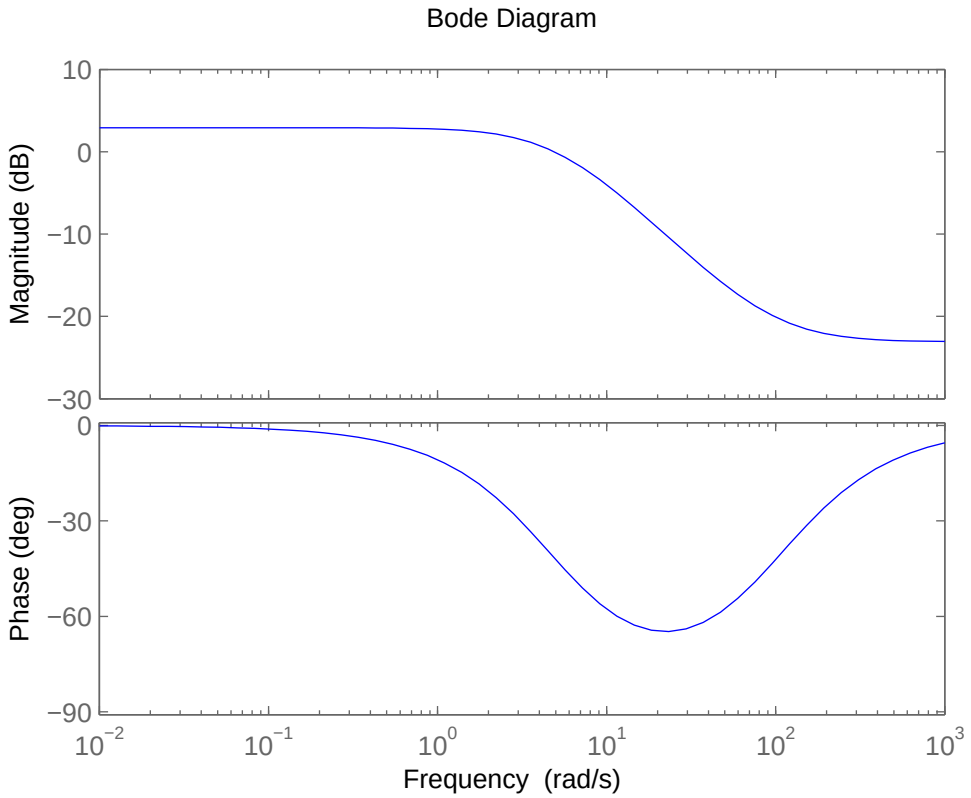

Figure 3.8: Bode plot of controller $C(s) = (0.07s + 7)/(s + 5)$

can be expressed as

$$|S(\mathrm{j}\omega)| < \frac{1}{|W_1(\mathrm{j}\omega)|} \tag{3.3.1}$$

$$|T(\mathrm{j}\omega)| < \frac{1}{|W_2(\mathrm{j}\omega)|} \tag{3.3.2}$$

where $W_1(s)$ and $W_2(s)$ are **weighting** transfer functions whose frequency responses satisfy

$$\frac{1}{|W_1(\mathrm{j}\omega)|} \ll 1 \quad \text{for } \omega < \omega_\mathrm{L} \tag{3.3.3}$$

$$\frac{1}{|W_2(\mathrm{j}\omega)|} \ll 1 \quad \text{for } \omega > \omega_\mathrm{H} \tag{3.3.4}$$

Since

$$|S(\mathrm{j}\omega)| = \frac{1}{|1 + G(\mathrm{j}\omega)C(\mathrm{j}\omega)|} \tag{3.3.5}$$

then making $|S(\mathrm{j}\omega)|$ for $\omega < \omega_\mathrm{L}$ means that $|G(\mathrm{j}\omega)C(\mathrm{j}\omega)|$ must be large over this frequency range, so that $|S(\mathrm{j}\omega)| \approx 1/|G(\mathrm{j}\omega)C(\mathrm{j}\omega)|$. This means that the

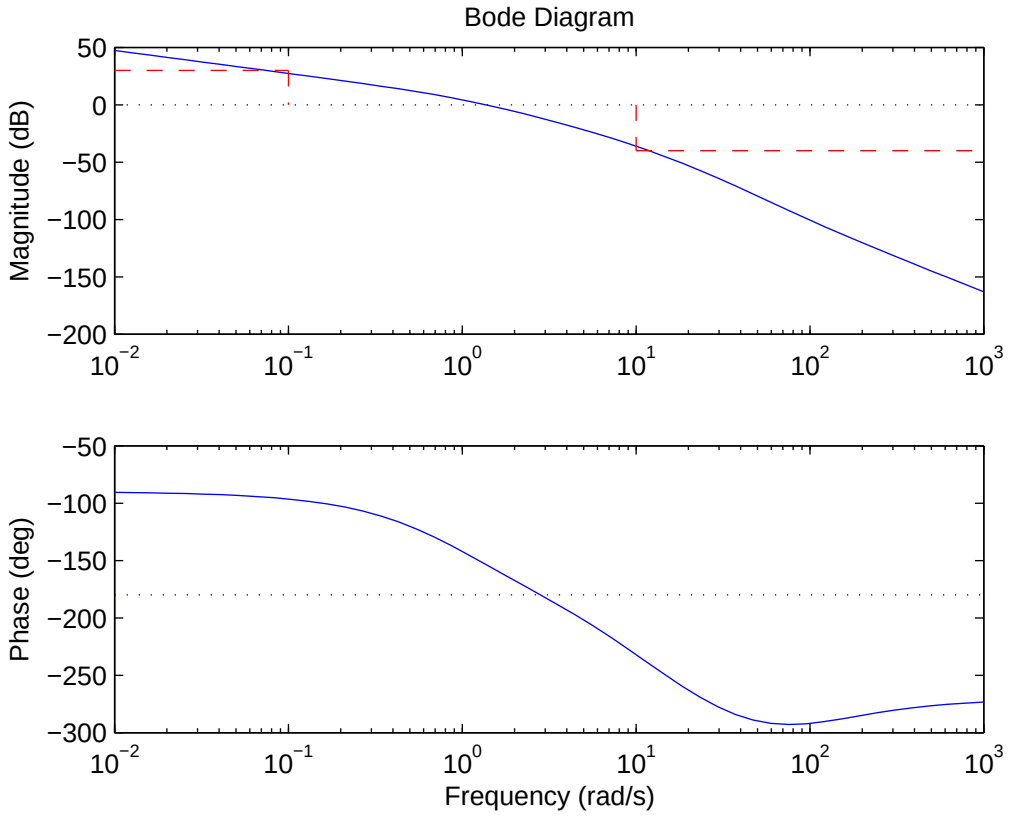

Figure 3.9: Bode plot of controller and system $G(s)C(s)$

condition in (3.3.1) is (approximately)

$$\frac{1}{|G(\mathrm{j}\omega)C(\mathrm{j}\omega)|} < \frac{1}{|W_1(\mathrm{j}\omega)|} \tag{3.3.6}$$

or

$$|G(\mathrm{j}\omega)C(\mathrm{j}\omega)| > |W_1(\mathrm{j}\omega)| \quad \text{for } \omega < \omega_\mathrm{L} \tag{3.3.7}$$

Similarly, since

$$|T(\mathrm{j}\omega)| = \frac{|G(\mathrm{j}\omega)C(\mathrm{j}\omega)|}{|1 + G(\mathrm{j}\omega)C(\mathrm{j}\omega)|} \tag{3.3.8}$$

then if $|T(\mathrm{j}\omega)|$ is small for $\omega > \omega_\mathrm{H}$, so that $|G(\mathrm{j}\omega)C(\mathrm{j}\omega)|$ must be small over this frequency range and the condition (3.3.2) becomes (approximately)

$$|G(\mathrm{j}\omega)C(\mathrm{j}\omega)| < \frac{1}{|W_2(\mathrm{j}\omega)|} \tag{3.3.9}$$

Hence, the conditions in (3.3.7) and (3.3.9) are in the same form as the conditions in (3.2.16) and (3.2.17), where the constants M_L and M_H have been replaced by a frequency dependent weighting.

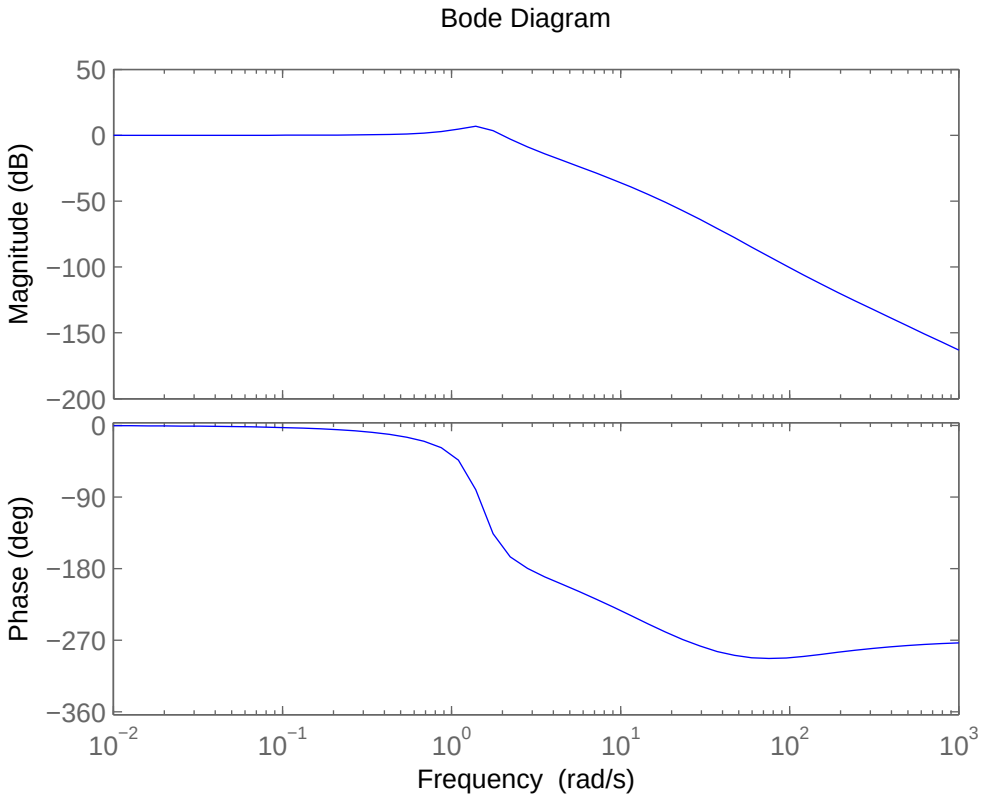

Figure 3.10: Bode plot of complementary sensitivity transfer function $T(s)$

Note. We can rearrange (3.3.1) and (3.3.2) as

$$|W_1(\mathrm{j}\omega)S(\mathrm{j}\omega)| < 1 \tag{3.3.10}$$

$$|W_2(\mathrm{j}\omega)T(\mathrm{j}\omega)| < 1 \tag{3.3.11}$$

and we will make use of this form in the next lecture.

Example 5 If the frequency weightings in the conditions in (3.3.7) and (3.3.9) are chosen to be

$$W_1(s) = \frac{0.56(s+0.5)}{s+00.5} \tag{3.3.12}$$

$$W_2(s) = \frac{5.62(s+5)}{s+50} \tag{3.3.13}$$

then the magnitude Bode plots of $W_1(\mathrm{j}\omega)$ (dotted red line) and $1/W_2(\mathrm{j}\omega)$ (dotted green line) are shown in Figure 3.12. Usually, the regions are defined in terms of the asymptotic Bode plots (solid lines) and these define the allowable regions for $|G(\mathrm{j}\omega)C(\mathrm{j}\omega)|$. The frequency response of an open loop system that satisfies these constraints is shown in blue. Using the frequency weightings to define the allowable regions “smooths” the corners of the regions.

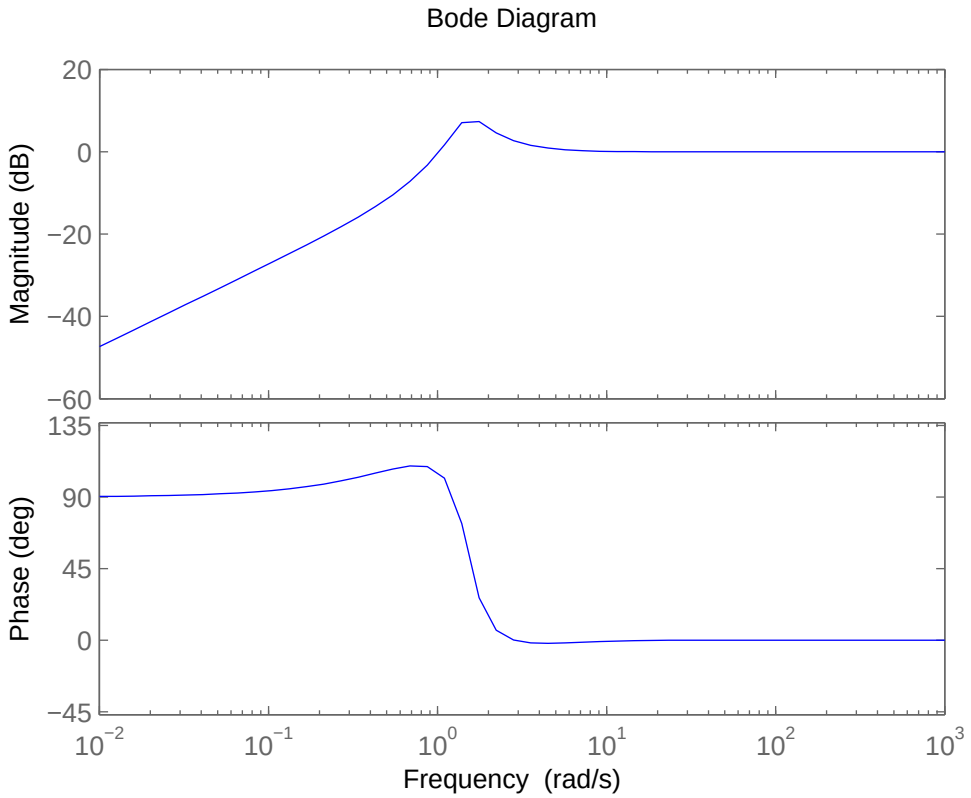

Figure 3.11: Bode plot of sensitivity transfer function $S(s)$

3.4 Bode Integral Theorem

If the magnitude Bode plot rolls-off at high frequencies faster than -40 dB per decade (which means that the order of the order of the denominator polynomial in $L(s)$ is at least two greater than the order of the numerator polynomial) and $\{p_k^{\mathrm{u}} : k = 1, 2, \ldots, M\}$ are the locations of the open loop **unstable** poles, then Bode's integral theorem gives a constraint on the frequency response of the sensitivity transfer function $S(s)$, which says that

$$\int_0^\infty \ln|S(\mathrm{j}\omega)|\mathrm{d}\omega = \pi \sum_{k=1}^{M} \mathrm{Re}\left(p_k^{\mathrm{u}}\right) \tag{3.4.1}$$

where M is the number of unstable open loop poles. If there are no unstable poles, so that $M = 0$, then

$$\int_0^\infty \ln|S(\mathrm{j}\omega)|\mathrm{d}\omega = 0 \tag{3.4.2}$$

This means that if we plot $\ln|S(\mathrm{j}\omega)|$ against ω, then the area of the plot under zero is the same as the area of the plot above zero. Because we are

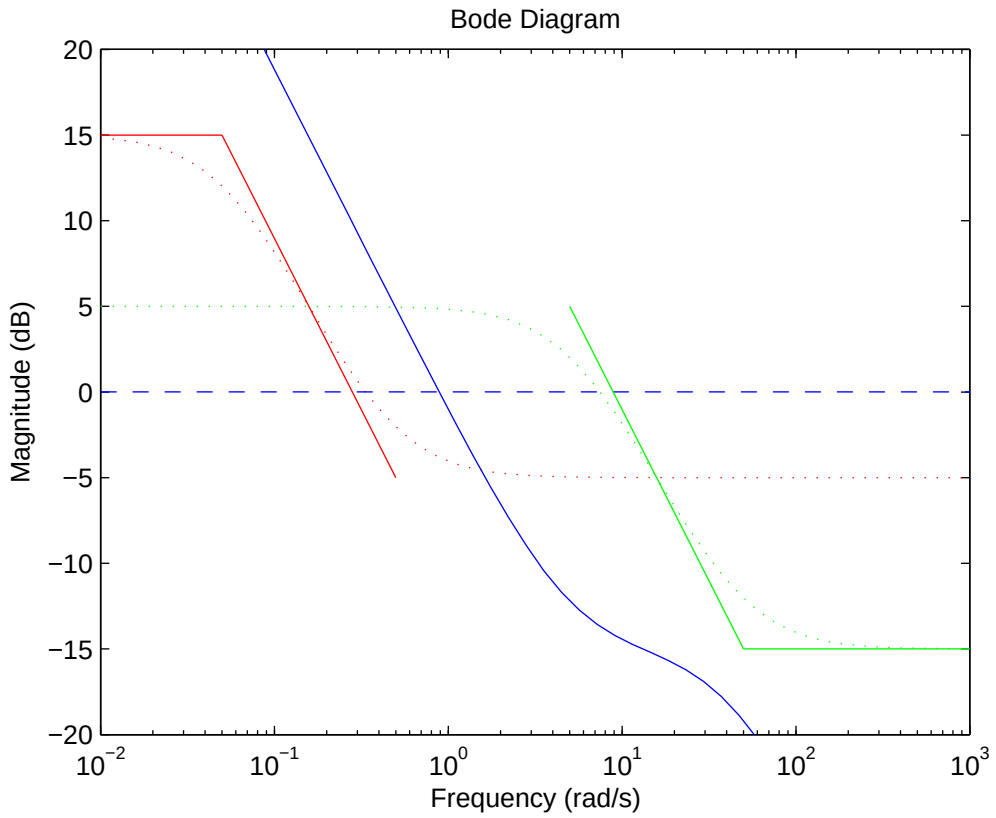

Figure 3.12: Magnitude Bode plots of $W_1(s)$ (red dotted) and $1/W_2(s)$ (green dotted) together with their asymptotic approximations (solid) and the frequency response of an open loop system (blue solid) that satisfies these constraints

designing the controller transfer function $C(s)$ with the aim of making $|S(\mathrm{j}\omega)|$ small for $\omega < \omega_c$, then $\ln|S(\mathrm{j}\omega)| < 0$ over this range of frequencies. However, in order to satisfy the Bode integral condition, there must be a range of frequencies above ω_c for which $\ln|S(\mathrm{j}\omega)|$ is greater than zero. As a result, $|S(\mathrm{j}\omega)| > 1$ in this region, meaning that the effect of the closed loop system is to **amplify** the effect of the disturbance on the output over this frequency range. This means that the effect of any components of the disturbance in the frequency range above ω_c will be amplified in the output. It is therefore important to ensure that ω_c is chosen so that it is above the bandwidth of $D(\mathrm{j}\omega)$, so that the magnitude of the components of the disturbance in this frequency range are negligible.

Example 6. Figure 3.13 shows $\ln|S(\mathrm{j}\omega)|$ plotted against ω. (Note that this is different from the Bode plot in Figure 3.11, which is plots $|S(\mathrm{j}\omega)|$ in dB against $\log_{10}\omega$, rather than plotting $\ln|S(\mathrm{j}\omega)|$ against ω). It can be seen that that the area below zero is balanced by an equal an equal area above zero.

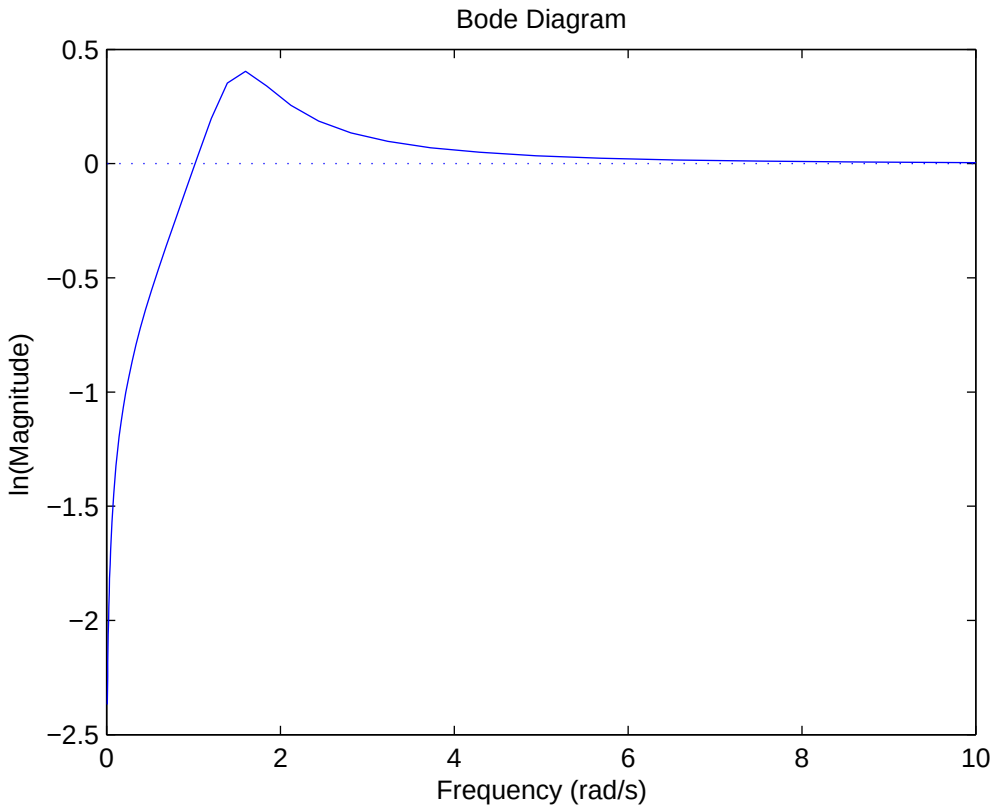

Figure 3.13: Plot of $\ln|S(j\omega)|$ against ω for system in Example 4

Bode's integral constraint is sometimes called the "waterbed" effect; if you push down $|S(j\omega)|$ for low frequencies, then this is offset by an increase in $|S(j\omega)|$ over higher frequencies. Further reducing $|S(j\omega)|$ at low frequencies and/or increasing ω_c, increases the area of $\ln|S(j\omega)|$ below zero, which causes a larger amplification of the disturbance at higher frequencies.

This discussion has considered the case where the open loop system is stable. If there are open loop unstable poles, then $\text{Re}\left(p_k^u\right)$ in (3.4.1) is greater than zero (since the open loop poles are in the right half plane), which means that the area of $\ln|S(j\omega)|$ above zero is larger than the area below zero. This makes the amplification of the high frequency components of the disturbance even worse, particularly if the poles are well to the right in the left half plane, so that $\sum_{k=1}^{M} \text{Re}\left(p_k^u\right)$ is large.

Note. The requirement that the Bode plot of $L(j\omega)$ rolls off at -40 dB per decade at high frequencies (included at the start of this section) ensures that there at least two more poles than zeros in $L(s)$, which means that $\lim_{s\to\infty} sL(s) = 0$. The more general form of the Bode integral which does

not include this restriction is

$$\int_0^\infty \ln|S(\mathrm{j}\omega)|\mathrm{d}\omega = \pi \sum_{k=1}^{M} \mathrm{Re}\left(p_k^{\mathrm{u}}\right) - \frac{\pi}{2} \lim_{s\to\infty} sL(s) \tag{3.4.3}$$

where M is the number of unstable open loop poles.

Lecture 4

Robust Control

Topics

1. Norms for signals and systems
2. Use of weighting transfer functions
3. Modelling uncertainty
4. Robustness and small gain theorem

Learning Outcomes

1. Understand how to determine the norms of signals and systems
2. Understand how a system can be manipulated so that it matches a standard setup
3. Understand how to express the uncertainty in a transfer function
4. Understand how to use the small gain theorem to guarantee that a closed loop system will remain stable in the presence of model uncertainty

4.1 Signal Norms

We need to be able to quantify the "size" of signals and we do this using norms. In the first lecture, we said that signals can be considered as functions of time, so for example, we can consider the input signal as a **mapping** from $t \in (-\infty, \infty)$ to the function $u(t)$. We define the $\mathcal{L}_p$ norm of this signal as

$$\|u(t)\|_p = \left(\int_{-\infty}^{\infty} |u(t)|^p \, \mathrm{d}t \right)^{\frac{1}{p}} \tag{4.1.1}$$

The most common norm that is used is the $\mathcal{L}_2$-norm, in which case

$$\|u(t)\|_2 = \left(\int_{-\infty}^{\infty} |u(t)|^2 \, \mathrm{d}t \right)^{\frac{1}{2}} \tag{4.1.2}$$

so that $\|u(t)\|_2^2$ is the total energy of the signal. If $\|u(t)\|_2 < \infty$, then we say that $u(t)$ is a finite energy signal. Often it is clear that we are using the $\mathcal{L}_2$ norm of the signal and we write $\|u(t)\|$ in place of $\|u(t)\|_2$

Taking Fourier transform of $u(t)$

$$U(\mathrm{j}\omega) = \int_{-\infty}^{\infty} u(t) \mathrm{e}^{-\mathrm{j}\omega t} \mathrm{d}t \tag{4.1.3}$$

then by Parseval's theorem

$$\begin{aligned} \|u(t)\|_2^2 &= \int_{-\infty}^{\infty} |u(t)|^2 \, \mathrm{d}t \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |U(\mathrm{j}\omega)|^2 \, \mathrm{d}\omega \end{aligned} \tag{4.1.4}$$

4.2 Norms of Systems

If we consider a system described by a transfer function $G(s)$, where

$$Y(s) = G(s)U(s) \tag{4.2.1}$$

then by taking the inverse Laplace transform, we know that

$$y(t) = \int_0^{\infty} g(\tau) u(t - \tau) \mathrm{d}\tau \tag{4.2.2}$$

where $g(t)$ is the impulse response. If $\int_0^\infty |g(\tau)|\,\mathrm{d}\tau < \infty$, then we showed that the system will be bounded-input, bounded-output (BIBO) stable. Taking the Fourier transform of $g(t)$

$$G(\mathrm{j}\omega) = \int_{-\infty}^{\infty} g(t)\mathrm{e}^{-\mathrm{j}\omega t}\mathrm{d}t \tag{4.2.3}$$

This is an alternative way of defining the frequency response, so that

$$Y(\mathrm{j}\omega) = G(\mathrm{j}\omega)U(\mathrm{j}\omega) \tag{4.2.4}$$

We can use this idea to define an (induced) norm for the system. For a system with transfer function $G(s)$, the $\mathcal{H}_\infty$ norm of the system is defined as

$$\begin{aligned} \|G\|_\infty &= \sup_{\|u(t)\|_2 \neq 0} \frac{\|Gu\|_2}{\|u\|_2} \\ &= \sup_{\|u(t)\|_2 \neq 0} \frac{\|y\|_2}{\|u\|_2} \end{aligned} \tag{4.2.5}$$

where $y(t)$ is the output obtained when the input $u(t)$ is applied to the system.

An alternative way of writing this is

$$\|G\|_\infty = \sup_\omega |G(\mathrm{j}\omega)| \tag{4.2.6}$$

which means that $\|G\|_\infty$ is the largest value of the magnitude of the frequency response of the system. To see this

$$\begin{aligned} \|y\|_2^2 &= \int_{-\infty}^{\infty} |y(t)|^2\,\mathrm{d}t \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |Y(\mathrm{j}\omega)|^2\,\mathrm{d}\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |G(\mathrm{j}\omega)U(\mathrm{j}\omega)|^2\,\mathrm{d}\omega \\ &\leq [\sup_\omega |G(\mathrm{j}\omega)|]^2 \frac{1}{2\pi}\int_{-\infty}^{\infty} |U(\mathrm{j}\omega)|^2\,\mathrm{d}\omega \end{aligned} \tag{4.2.7}$$

so that

$$\|y\|_2 \leq [\sup_\omega |G(\mathrm{j}\omega)|]\,\|u\|_2 \tag{4.2.8}$$

and

$$\begin{aligned} \frac{\|y\|_2}{\|u\|_2} &= \frac{\|Gu\|_2}{\|u\|_2} \\ &\leq \sup_\omega |G(\mathrm{j}\omega)| \end{aligned} \tag{4.2.9}$$

This means that the **gain** of the system from input to output is at most $\sup_\omega |G(j\omega)|$. We call this gain the $\mathcal{H}_\infty$ norm of the system. The $\mathcal{H}_\infty$ norm is the largest value of the magnitude of the frequency response, which as shown in From Figure 4.1, can be read directly from the magnitude Bode plot.

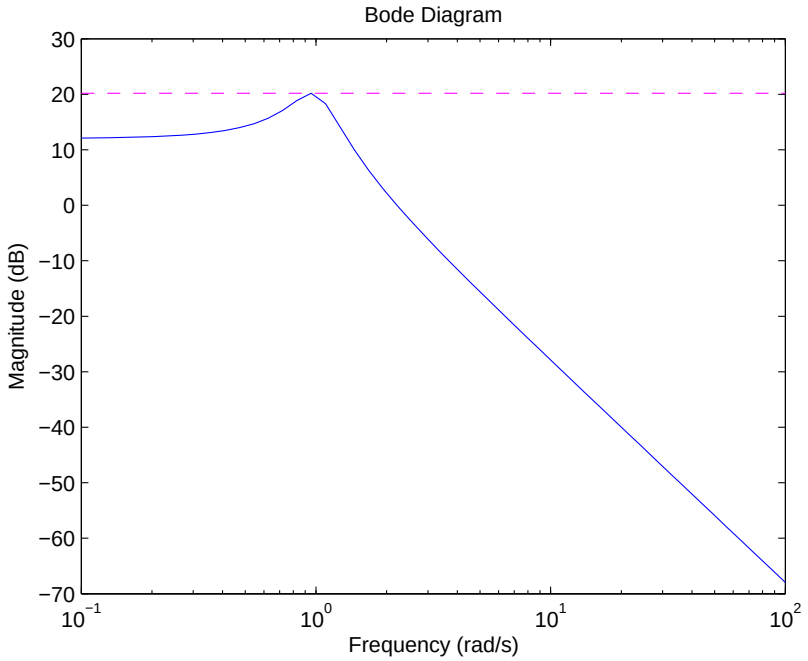

Figure 4.1: Magnitude frequency response $|G(j\omega)|$ together with maximum value

Note 1. The $\mathcal{H}_\infty$ norm is defined in terms of the maximum ratio of the $\mathcal{L}_2$ norms of the signals (not the $\mathcal{L}_\infty$ norms of the signals).

Note 2. The norm is defined on a Hardy space consisting of rational transfer functions that describe **causal** systems.

4.3 Standard Setup

So far we have been considering the arrangement of the feedback system as shown in Figure 4.2. However, we can rearrange the system to the form in Figure 4.3 where we have grouped together the “exogenous” input signals

(i.e. the signals are generated outside the system) into a vector $\mathbf{w}(t)$

$$\mathbf{w}(t) = \begin{bmatrix} n(t) \\ d_o(t) \\ d_i(t) \\ r(t) \end{bmatrix} \tag{4.3.1}$$

Similarly $\mathbf{z}(t)$ collects together the signals that we wish to control

$$\mathbf{z}(t) = \begin{bmatrix} z(t) \\ u(t) \\ e(t) \end{bmatrix} \tag{4.3.2}$$

This is a very general arrangement of the closed loop system; often not all of the signals will be considered. We can redraw the block diagram as in Figure 4.4, where the controller has been included in the feedback loop.

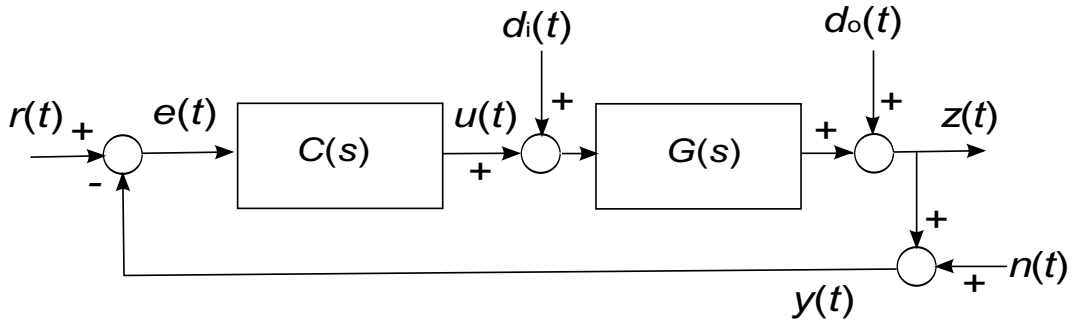

Figure 4.2: Signals in closed loop

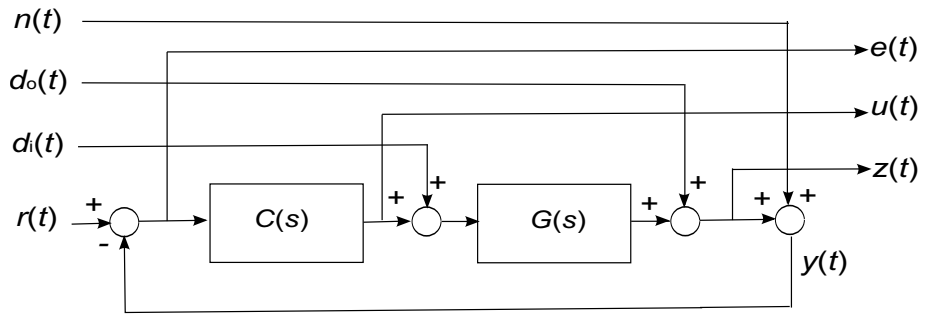

Figure 4.3: Signals in closed loop with exogenous inputs and outputs grouped together

Example. LQG Control For a system described by a state space model

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) + \mathbf{d}_i(t) \\ y(t) &= \mathbf{c}^T\mathbf{x}(t) + n(t) \end{aligned} \tag{4.3.3}$$

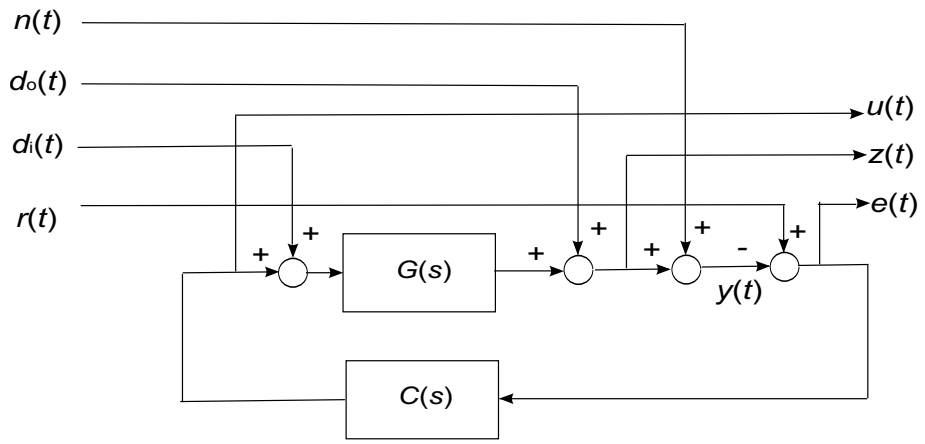

Figure 4.4: Signals in closed loop with controller in feedback loop

where $\mathbf{d}_\mathrm{i}(t)$ is zero-mean, white noise (vector) process noise, with variance $E\left[\mathbf{d}_\mathrm{i}(t)\mathbf{d}_\mathrm{i}^\mathsf{T}(t)\right] = \mathbf{Q}_\mathrm{d}$ and $n(t)$ is zero-mean, white noise sensor noise, with variance $E\left[n^2(t)\right] = r_\mathrm{d}$. The LQG controller is designed to minimise the cost function

$$J = \int_0^\infty \mathbf{x}(t)^\mathsf{T}\mathbf{Q}_\mathrm{c}\mathbf{x}(t) + r_\mathrm{c}u^2(t)\mathrm{d}t \tag{4.3.4}$$

where $\mathbf{Q}_\mathrm{c} \in \mathbb{R}^{n\times n}$ is a positive matrix (real part of all eigenvalues greater or equal to zero) and $r_\mathrm{c} > 0$, which determine the relative contributions of the states and the input to the cost function. Define

$$\mathbf{w}(t) = \begin{bmatrix} \mathbf{d}_i(t) \\ n(t) \end{bmatrix} \tag{4.3.5}$$

and

$$\mathbf{z}(t) = \begin{bmatrix} \mathbf{x}(t) \\ u(t) \end{bmatrix} \tag{4.3.6}$$

as shown in Figure 4.5.

The cost function can be written as

$$\begin{aligned} J &= \int_0^\infty \mathbf{x}(t)^\mathsf{T}\mathbf{Q}_\mathrm{c}\mathbf{x}(t) + r_\mathrm{c}u^2(t)\mathrm{d}t \\ &= \int_0^\infty \mathbf{z}(t)^\mathsf{T}\begin{bmatrix} \mathbf{Q}_\mathrm{c} & 0 \\ 0 & r_\mathrm{c} \end{bmatrix}\mathbf{z}(t)\mathrm{d}t \end{aligned} \tag{4.3.7}$$

The controller determines the input $u(t)$ that minimises this cost function.

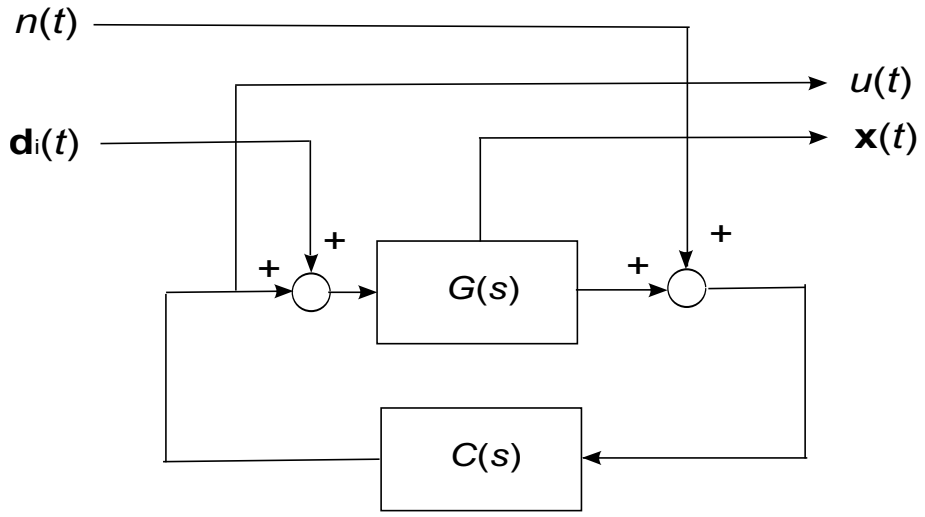

Figure 4.5: Signals for LQG controller

From the B15 lectures, we know that the optimal controller takes the form of a state feedback

$$u(t) = -\mathbf{f}^{\mathsf{T}}\hat{\mathbf{x}}(t) \tag{4.3.8}$$

where $\mathbf{f} \in \mathbb{R}^n$ is the state feedback matrix (which is a vector in this case since we are considering a SISO system) and $\hat{\mathbf{x}}(t)$ is the estimate of the state generated by the Kalman filter, which is the optimal observer

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{b}u(t) + \mathbf{k}\left[y(t) - \mathbf{c}^{\mathsf{T}}\hat{\mathbf{x}}(t)\right] \tag{4.3.9}$$

where $\mathbf{k} \in \mathbb{R}^n$ is the Kalman gain. The optimal state feedback is given by

$$\mathbf{f} = r_{\mathrm{c}}^{-1}\mathbf{b}^{\mathsf{T}}\mathbf{P}_{\mathrm{f}} \tag{4.3.10}$$

where $\mathbf{P}_{\mathrm{f}} \in \mathbb{R}^{n\times n}$ is the solution to the algebraic Riccati equation

$$\mathbf{P}_{\mathrm{f}}\mathbf{A} + \mathbf{A}^{\mathsf{T}}\mathbf{P}_{\mathrm{f}} - \mathbf{P}_{\mathrm{f}}\mathbf{b}r_{\mathrm{c}}^{-1}\mathbf{b}^{\mathsf{T}}\mathbf{P}_{\mathrm{f}} + \mathbf{Q}_{\mathrm{c}} = \mathbf{0} \tag{4.3.11}$$

Similarly, the optimal Kalman gain is

$$\mathbf{k} = \mathbf{P}_{\mathrm{o}}\mathbf{c}r_{\mathrm{d}}^{-1} \tag{4.3.12}$$

where $\mathbf{P}_{\mathrm{o}} \in \mathbb{R}^{n\times n}$ is the solution to the algebraic Riccati equation

$$\mathbf{P}_{\mathrm{o}}\mathbf{A} + \mathbf{A}^{\mathsf{T}}\mathbf{P}_{\mathrm{o}} - \mathbf{P}_{\mathrm{o}}\mathbf{c}r_{\mathrm{d}}^{-1}\mathbf{c}^{\mathsf{T}}\mathbf{P}_{\mathrm{o}} + \mathbf{Q}_{\mathrm{d}} = \mathbf{0} \tag{4.3.13}$$

Combining (4.3.8) and (4.3.9)

$$\dot{\hat{\mathbf{x}}}(t) = \left[\mathbf{A} - \mathbf{b}\mathbf{f}^{\mathsf{T}} - \mathbf{k}\mathbf{c}^{\mathsf{T}}\right]\hat{\mathbf{x}}(t) + \mathbf{k}y(t) \tag{4.3.14}$$

$$u(t) = -\mathbf{f}^{\mathsf{T}}\hat{\mathbf{x}}(t) \tag{4.3.15}$$

This is the state space model of the optimal controller which takes $y(t)$ as the input to the controller and generates $u(t)$ as the output from the controller. The model can be written more compactly as

$$\begin{bmatrix} \dot{\hat{\mathbf{x}}}(t) \\ u(t) \end{bmatrix} = \left[\begin{array}{c|c} \mathbf{A} - \mathbf{b}\mathbf{f}^{\mathsf{T}} - \mathbf{k}\mathbf{c}^{\mathsf{T}} & \mathbf{k} \\ \hline -\mathbf{f}^{\mathsf{T}} & 0 \end{array}\right] \begin{bmatrix} \hat{\mathbf{x}}(t) \\ y(t) \end{bmatrix} \tag{4.3.16}$$

4.4 Weightings

For the LQG controller, cost function in (4.3.4) depends upon $\mathbf{x}^{\mathsf{T}}(t)\mathbf{Q}_{\mathrm{c}}\mathbf{x}(t)$ and $r_{\mathrm{c}}u^2(t)$. If we define

$$\tilde{u}(t) = w_u u(t) \tag{4.4.1}$$

where $w_u^2 = r_{\mathrm{c}}$ and

$$\tilde{\mathbf{x}}(t) = \mathbf{W}_x \mathbf{x}(t) \tag{4.4.2}$$

where $\mathbf{W}_x \in \mathbb{R}^{n\times n}$ satisfies $\mathbf{W}_x^{\mathsf{T}}\mathbf{W}_x = \mathbf{Q}_{\mathrm{c}}$, then the cost function becomes

$$J = \int_0^\infty \tilde{\mathbf{x}}(t)^{\mathsf{T}}\tilde{\mathbf{x}}(t) + \tilde{u}^2(t)\mathrm{d}t \tag{4.4.3}$$

The exogenous inputs are $\mathbf{d}_{\mathrm{i}}(t)$ and $n(t)$. These are both zero-mean, white noise signals with variances $\mathbf{Q}_{\mathrm{d}}$ and r_{d} respectively. If we define $\tilde{n}(t)$ such that

$$n(t) = w_n \tilde{n}(t) \tag{4.4.4}$$

where $w_n^2 = r_{\mathrm{d}}$, then $\tilde{n}(t)$ is a zero-mean, white noise process that has **unit** variance.

Similarly, if we define $\tilde{\mathbf{d}}_{\mathrm{i}}(t) \in \mathbb{R}^n$ such that

$$\mathbf{d}_{\mathrm{i}}(t) = \mathbf{W}_d \tilde{\mathbf{d}}_{\mathrm{i}}(t) \tag{4.4.5}$$

where $\mathbf{W}_d \in \mathbb{R}^{n\times n}$ satisfies $\mathbf{W}_d^{\mathsf{T}}\mathbf{W}_d = \mathbf{Q}_{\mathrm{d}}$, then $\tilde{\mathbf{d}}_{\mathrm{i}}(t)$ is zero-mean, white noise process whose covariance matrix is the identity matrix.

We can now modify the feedback loop to the form shown in Figure 4.6. In this form, the exogenous inputs are zero-mean, white noise processes that

have unit variance. In practice, it is unrealistic to assume that the exogenous signals will be white noise, as real disturbances are unlikely to have a flat spectrum. As we saw in Lecture 3, it is likely that most of the components of the reference signal and the disturbances will be in the low frequency range, while the sensor noise is likely to have high frequency components. If we allow $\mathbf{W}_d$ and w_n to be transfer functions, so that

$$D_{\mathrm{i}}(\mathrm{j}\omega) = \mathbf{W}_d(\mathrm{j}\omega)\tilde{D}_{\mathrm{i}}(\mathrm{j}\omega) \tag{4.4.6}$$

$$N(\mathrm{j}\omega) = W_n(\mathrm{j}\omega)\tilde{N}(\mathrm{j}\omega) \tag{4.4.7}$$

then by appropriate choice of $\mathbf{W}_d(\mathrm{j}\omega)$ and $W_n(\mathrm{j}\omega)$ the expected frequency content of the exogenous signals can be modelled.

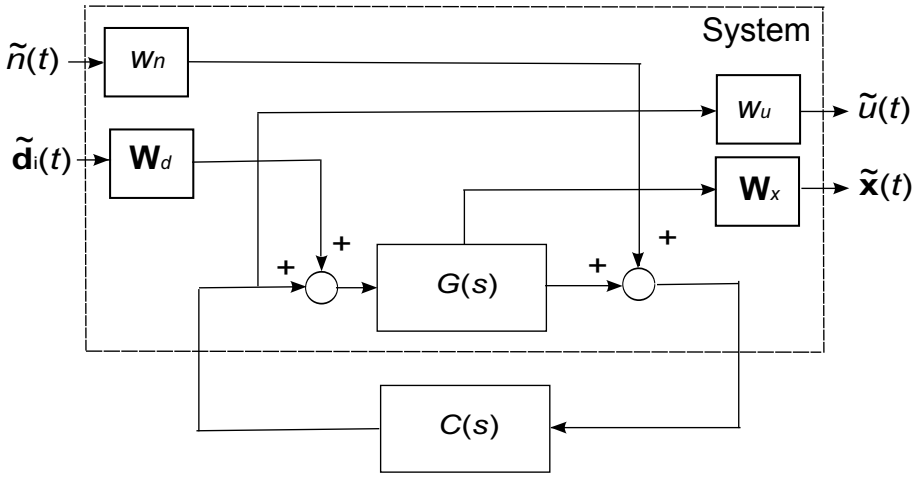

Figure 4.6: Weighted signals for LQG controller

We can also include frequency dependent weightings on the output. For example, from Lecture 3, we now that we want the mapping from the disturbance to the output, defined by the frequency response of sensitivity function $|S(\mathrm{j}\omega)|$, to be small at low frequencies, while the frequency response of the complementary sensitivity function $|T(\mathrm{j}\omega)|$ needs to be small at high frequencies.

If we incorporate all of the weightings into the system model, then we can arrive at the description in 4.7, where $\tilde{G}(s)$ is the **combined** transfer function of $G(s)$ and the weighting functions. This can be written compactly in the

form

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{z}(t) \\ y(t) \end{bmatrix} = \left[\begin{array}{c|cc} \mathbf{A} & \mathbf{B}_1 & \mathbf{b}_2 \\ \hline \mathbf{C}_1 & \mathbf{D}_{11} & \mathbf{d}_{12} \\ \mathbf{c}_2^{\mathsf{T}} & \mathbf{d}_{21}^{\mathsf{T}} & d_{22} \end{array}\right] \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{w}(t) \\ u(t) \end{bmatrix} \tag{4.4.8}$$

Note that in LQG control, $\mathbf{D}_{11}$ and d_{22} are both zero. If we design an LQG controller for a system that is augmented by weighting functions, then this is referred to as $\mathcal{H}_2$ control.

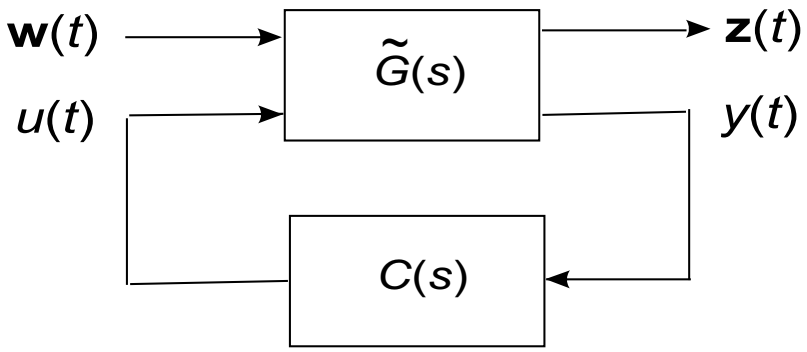

Figure 4.7: General form of controller

4.5 Uncertainty

The design of a control system is based upon the transfer function of the system, $G(s)$. In practice, we do not know $G(s)$ exactly due to uncertainties in the modelling that generates the transfer function. These uncertainties can come from unmodelled dynamics or the uncertainties associated with the estimation of the parameters. There are a number of ways of expressing the uncertainty, but one way is to consider the uncertainty as **additive**. We call the transfer function associated with our ideal model of the system the **nominal** transfer function, which we denote by $G_0(s)$. The actual transfer function will be $G(s)$.

We do not know exactly what $G(s)$ is, but we know that it is close (in some sense) to $G_0(s)$. For example, we could model the uncertainty by

$$|G(\mathrm{j}\omega) - G_0(\mathrm{j}\omega)| \leq 0.1 \quad \text{for all } \omega \tag{4.5.1}$$

which gives a bound on the distance between the nominal frequency response $G_0(j\omega)$ and the actual frequency response $G(j\omega)$. Using the definition of the $\mathcal{H}_\infty$ norm in (4.2.6)

$$\begin{aligned} \|G(j\omega) - G_0(j\omega)\|_\infty &= \sup_\omega |G(j\omega) - G_0(j\omega)| \\ &\leq 0.1 \end{aligned} \tag{4.5.2}$$

We can turn this round to say

$$G(s) = G_0(s) + 0.1\Delta(s) \tag{4.5.3}$$

where $\Delta(s)$ is an unknown, stable transfer function that satisfies

$$\|\Delta\|_\infty \leq 1 \tag{4.5.4}$$

This analysis has considered the case where the bound on the magnitude of the difference between $G_0(j\omega)$, the nominal frequency response, and $G(j\omega)$ is constant for all frequencies. In practice, the bound is likely to depend upon frequency; for example, the model tends to be accurate at low frequencies, so the magnitude of the difference is small at these frequencies. Because the system must be strictly proper, at high frequencies, the magnitude of $G_0(s)$ and $G(s)$ tends to be small, so the difference between them will also be small. At frequencies between these two extremes, the magnitude of the difference could be relatively large. The frequency dependence of the difference can be modelled using

$$|G(j\omega) - G_0(j\omega)| \leq |W_\Delta(j\omega)| \tag{4.5.5}$$

so that we can write

$$G(s) = G_0(s) + W_\Delta(s)\Delta(s) \tag{4.5.6}$$

where $W_\Delta(s)$ is weighting transfer function and $\Delta(s)$ satisfies the constraint in (4.5.4).

The uncertainty can be incorporated into the system model as shown in Figure 4.8. If the weighting $W_\Delta(s)$ is combined with $G(s)$ and the other weightings into $\tilde{G}_0(s)$, then the system can be described compactly as in Figure 4.9.

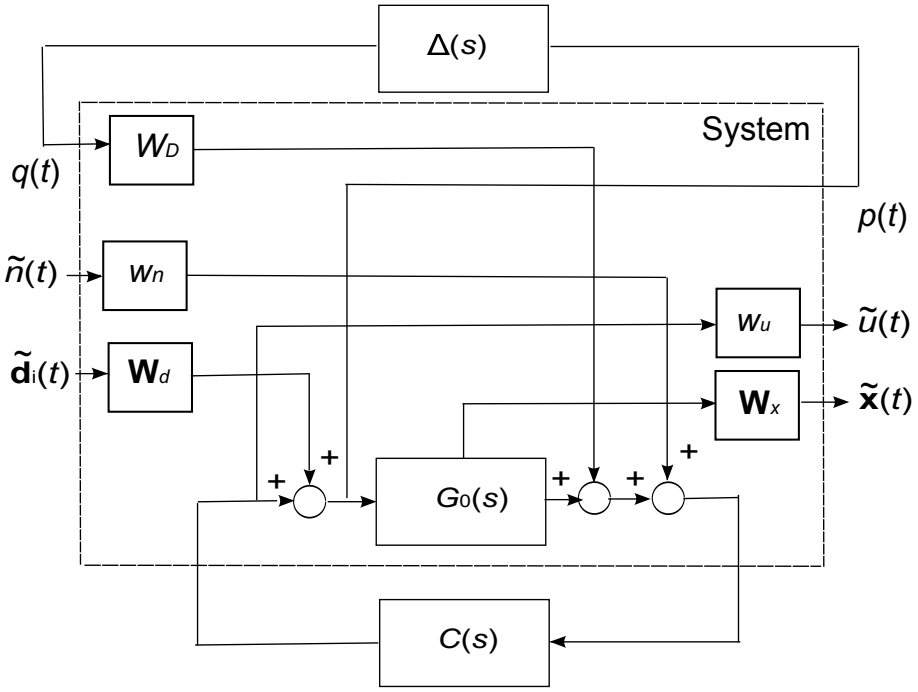

Figure 4.8: System in Figure 4.6 including uncertainty

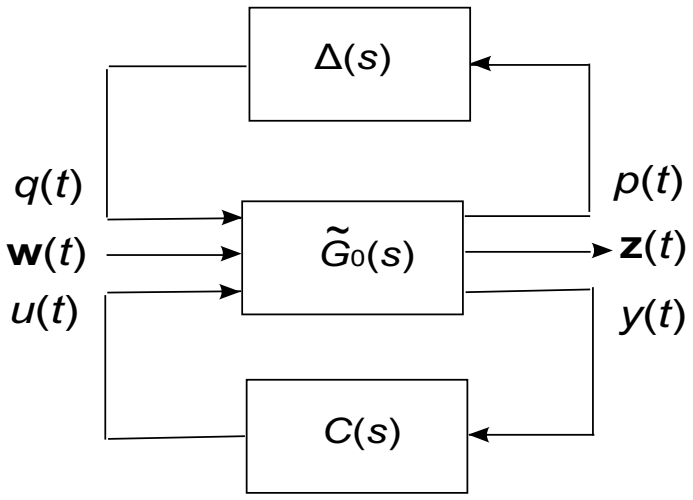

Figure 4.9: Compact form of system including uncertainty

The other common form of uncertainty description is the multiplicative form

$$G(s) = G_0(s)\,(1 + W_\Delta(s)\Delta(s)) \tag{4.5.7}$$

The additive form is expressed in terms of the absolute magnitude of the difference between $G_0(\mathrm{j}\omega)$ and $G(\mathrm{j}\omega)$, while the multiplicative form considers the difference as a fraction of the magnitude of $G_0(\mathrm{j}\omega)$. The structure of the two forms is shown in Figure 4.10

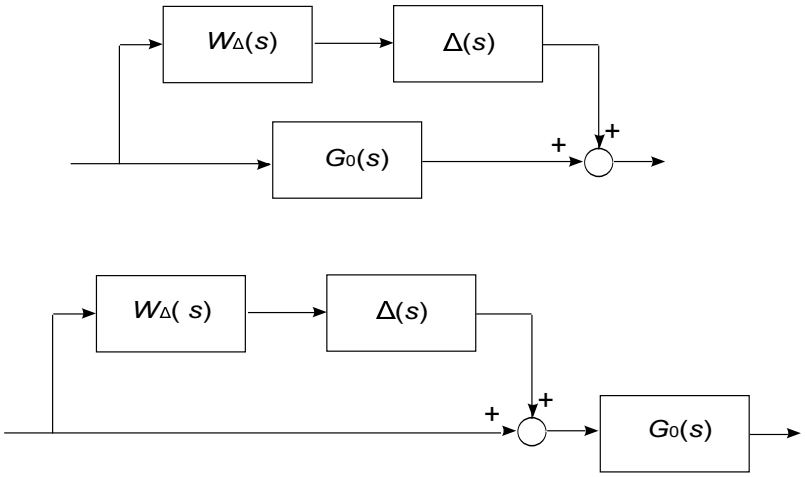

Figure 4.10: Structure of additive uncertainty (top) and multiplicative uncertainty (bottom)

4.6 Robustness to Model Uncertainty

We design the controller $C(s)$ on the basis of the transfer function of the nominal system $G_0(s)$ and as part of the design, we ensure that the closed loop system is stable when controlling $G_0(s)$. The problem is that if there model uncertainties, the actual system may have a different (unknown) transfer function $G(s)$ and it is possible that the closed loop system for the actual transfer function may be unstable. To test whether the feedback system will remain stable for the unknown transfer function that incorporates the uncertainty, we make use of the following theorem

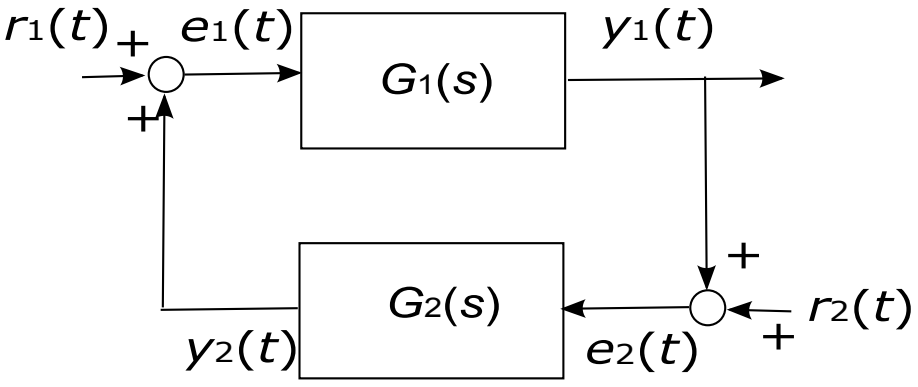

Figure 4.11: Feedback system for proof of Small Gain Theorem

Small Gain Theorem The closed loop system in Figure 4.11 is stable provided

$$\|G_1(s)\|_\infty \|G_2(s)\|_\infty < 1 \tag{4.6.1}$$

Note that in the diagram, the signals are **added** at the summing junctions, but the theorem would still hold, even if one of signals are subtracted at a junction, for example, as in the case for negative feedback.

Outline Proof. A full proof requires some subtle mathematical arguments about extended spaces, so we give an "informal" proof. Suppose that $r_1(t)$ and $r_2(t)$ are finite energy signals, so that

$$\|r_1\|_2 < \infty \quad \text{and } \|r_2\|_2 < \infty \tag{4.6.2}$$

then from Figure 4.11

$$\begin{aligned} e_1 &= r_1 + y_2 && (4.6.3)\\ &= r_1 + G_2 e_2 && (4.6.4)\\ &= r_1 + G_2 r_2 + G_2 G_1 e_1 && (4.6.5) \end{aligned}$$

so that

$$\begin{aligned} \|e_1\|_2 &\leq \|r_1\|_2 + \|G_2 r_2\|_2 + \|G_2 G_1 e_1\|_2 && (4.6.6)\\ &\leq \|r_1\|_2 + \|G_2\|_\infty \|r_2\|_2 + \|G_2 G_1\|_\infty \|e_1\|_2 && (4.6.7)\\ &\leq \|r_1\|_2 + \|G_2\|_\infty \|r_2\|_2 + \|G_1\|_\infty \|G_2\|_\infty \|e_1\|_2 && (4.6.8)\\ &\leq \frac{\|r_1\|_2 + \|G_2\|_\infty \|r_2\|_2}{1 - \|G_1\|_\infty \|G_2\|_\infty} && (4.6.9) \end{aligned}$$

If $\|G_1\|_\infty \|G_2\|_\infty < 1$, then $\|e_1\|_2 < \infty$, and if $\|r_1\|_2 < \infty$ and $\|e_1\|_2 < \infty$, then $\|y_2\|_2 < \infty$. By a similar argument, it is possible to show that $\|e_2\|_2 < \infty$ and $\|y_1\|_2 < \infty$. Hence the condition

$$\|G_1\|_\infty \|G_2\|_\infty < 1 \tag{4.6.10}$$

is sufficient for stability.

4.6.1 Robustness to Multiplicative Uncertainty

Suppose that we have a system that is subject to a multiplicative uncertainty $W_\Delta(s)\Delta(s)$, as in Figure 4.10, where the weighting is chosen so that $\|\Delta\|_\infty < 1$. When this system forms part of a feedback loop with a controller $C(s)$

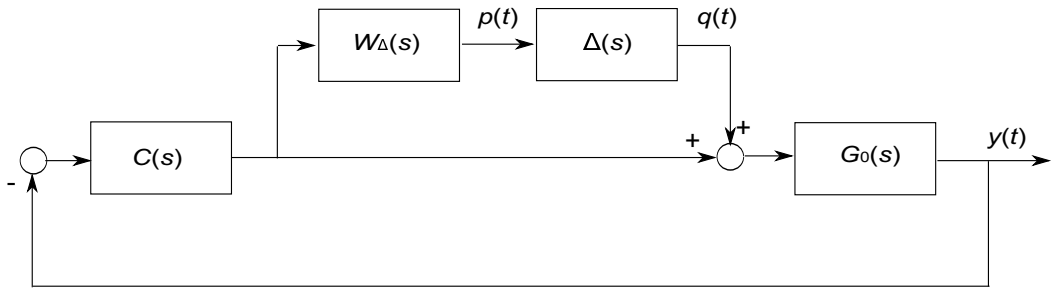

Figure 4.12: System with multiplicative uncertainty connected in feedback loop

shown in Figure 4.12, then the closed loop transfer function including the uncertainty is

$$Q(s) = \Delta(s)P(s) \tag{4.6.11}$$

(4.6.12)

$$Y(s) = G(s)Q(s) - G(s)C(s)Y(s) \tag{4.6.13}$$

$$= \frac{G(s)}{1 + G(s)C(s)}Q(s) \tag{4.6.14}$$

$$P(s) = -W_\Delta(s)C(s)Y(s) \tag{4.6.15}$$

$$= -W_\Delta(s)\frac{C(s)G(s)}{1 + G(s)C(s)}Q(s) \tag{4.6.16}$$

$$= -W_\Delta(s)T(s)Q(s) \tag{4.6.17}$$

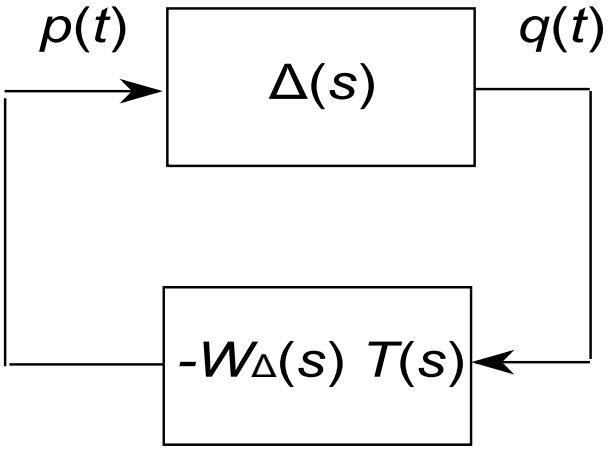

Figure 4.13: Alternative representation of feedback loop in Figure 4.12

The feedback loop can therefore be expressed in the form in Figure 4.13 and using the Small Gain Theorem, the system will be stable for any uncertainty satisfying $\|\Delta\|_\infty < 1$, provided that

$$\|W_\Delta T\|_\infty < 1 \tag{4.6.18}$$

Since

$$\|W_\Delta T\|_\infty = \sup_\omega |W_\Delta(j\omega)T(j\omega)| \tag{4.6.19}$$

then a sufficient condition to satisfy (4.6.18) is

$$|W_\Delta(j\omega)T(j\omega)| < 1 \quad \text{for all } \omega \tag{4.6.20}$$

Using

$$|W_\Delta(j\omega)T(j\omega)| = \left| W_\Delta(j\omega)\frac{L(j\omega)}{1+L(j\omega)} \right| \tag{4.6.21}$$

where $L(s) = G(s)C(s)$, the robust stability condition will be satisfied provided

$$|W_\Delta(j\omega)L(j\omega)| < |1+L(j\omega)| \tag{4.6.22}$$

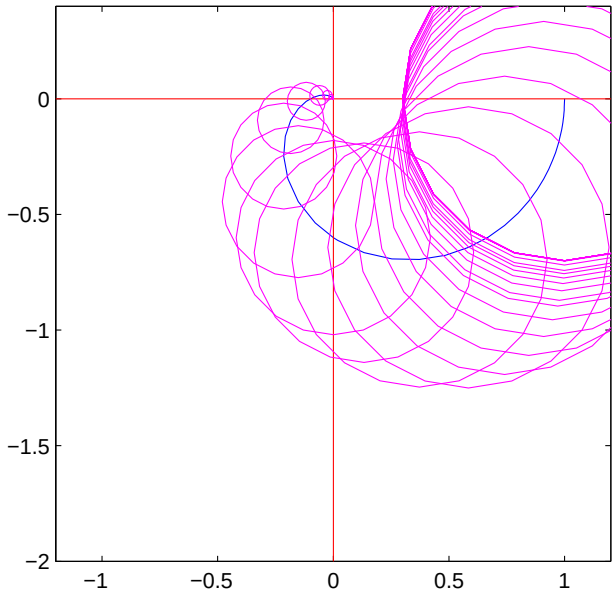

Figure 4.14: Nyquist plot for system with multiplicative uncertainty

This condition can be “tested” graphically using the Nyquist diagram as shown in Figure 4.14. For each frequency ω, the right hand side of (4.6.22)

Perturbation	$\tilde{G}(s)$	Condition
Multiplicative	$(1 + \Delta(s)W(s))\,G(s)$	$\|\|W(s)T(s)\|\|_\infty < 1$
Additive	$G(s) + \Delta(s)W(s)$	$\|\|W(s)C(s)S(s)\|\|_\infty < 1$
Inverse Multiplicative	$G(s)/\,(1 + \Delta(s)W(s)G(s))$	$\|\|W(s)G(s)S(s)\|\|_\infty < 1$
Inverse Additive	$G(s)/\,(1 + \Delta(s)W(s))$	$\|\|W(s)S(s)\|\|_\infty < 1$

Table 4.1: Conditions for robust stability for different forms of uncertainty

$|1 + L(\mathrm{j}\omega)|$, is the distance from the point $-1 + \mathrm{j}0$ to $L(\mathrm{j}\omega)$, while $|W_\Delta(\mathrm{j}\omega)L(\mathrm{j}\omega)|$ is the distance from the origin to $L(\mathrm{j}\omega)$, scaled by $W_\Delta(\mathrm{j}\omega)$. This means that if a circle of radius $|W_\Delta(\mathrm{j}\omega)L(\mathrm{j}\omega)|$ centred on each point on the Nyquist plot does not encircle $-1 + \mathrm{j}0$ then the closed loop system is stable for any uncertainty that satisfies $\|\Delta\|_\infty < 1$.

4.6.2 Robustness to Other Forms of Uncertainty

The analysis in the preceding section assumed that the uncertainty was in the form of a multiplicative perturbation to a nominal model. However, the small gain theorem can be used to determine the stability of the closed loop system in the presence of other forms of uncertainty, as set out in Table 4.1

The small gain theorem is used here to **analyse** whether a given controller $C(s)$ is stable when connected in a feedback loop with an uncertain system. It is a much more difficult problem to **design** a controller that will be guarantee robust stability and also whether the performance of the closed loop system will still satisfy the control specification - this will be covered in the C20 courses.