1 EL2520 Lecture notes 1: Introduction, Norms of Signals and Systems, The Small Gain Theorem

Feedback is a relatively simple yet extremely powerful principle for modifying the behavior of dynamical systems. Feedback is everywhere, both in natural/biological and engineered systems. The basic principle is that decisions are based on observations of the system behavior rather than on prior knowledge and information about external influences. The main reason for using feedback is that it is the most effective way to deal with uncertainty about a system and its surroundings. This course presents methods for analysing and designing feedback control algorithms for multivariable systems with respect to stability and performance in the presence of uncertainty.

The standard feedback control problem can be represented by the block diagram in Figure 1.1. Here G represents the system to be controlled, F_y is the feedback controller and F_r is a pre-filter on the setpoint r. The other signals (variables) are the plant input (manipulated variable) u, the plant output z, the measurement y, the disturbance on the output w and the measurement noise n. The aim of the control system is to make the output z follow the setpoint r in the presence of disturbances w and measurement noise n, or equivalently, keep the control error $e = r - z$ small[1]. It is usually of interest to also keep the control input u small in some sense.

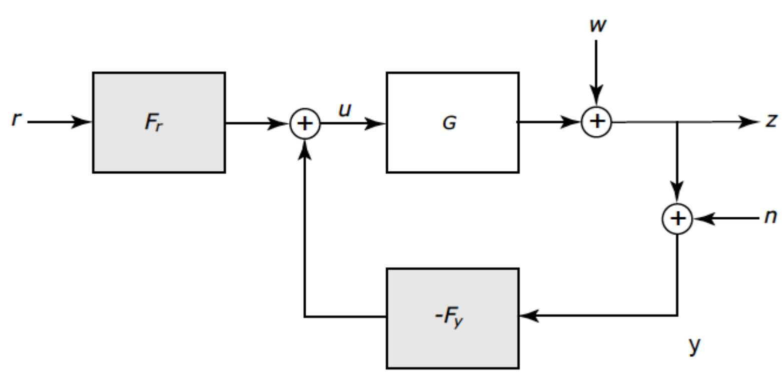

Figure 1.1: Two degree of freedom control system.

The feedback controller in Figure 1.1 is a so-called two degree of freedom controller since we have a prefilter F_r on the setpoint and a feedback controller F_y based on the measurement y

$$u = F_r r - F_y y$$

[1]The formulation of the control problem as a setpoint tracking problem may seem limiting, but note that many problems can be translated into this form. For instance, (local) optimization of a system can be achieved by letting the output be the derivative of the objective function and the setpoint $r = 0$.

The two degrees of freedom allow us to design the response to setpoints r more or less independently of the response to disturbances w and noise n. The standard one degree of freedom controller is obtained by letting $F_r = F_y = F$ which yields $u = F(r - y)$.

In this course we will consider multivariable systems, i.e., all signals are vectors and all transfer-functions are matrices. For instance, consider a 2×2 system with two inputs u_1, u_2 and two outputs y_1, y_2. We then have $u = [u_1 \; u_2]^T$ and $y = [y_1 \; y_2]^T$ and the transfer matrix $G(s)$ is a 2×2 matrix with the element $G_{ij}(s)$ being the transfer-function from input u_j to output y_i.

The focus will be on systems that can be described by a linear time invariant (LTI) model, which on state space form is

$$\begin{aligned} \dot{x} &= Ax(t) + Bu(t) \; ; x \in \mathbb{R}^n, u \in \mathbb{R}^m \\ y(t) &= Cx(t) + Du(t) \; ; y \in \mathbb{R}^p \end{aligned} \tag{1}$$

where x is the n-dimensional state vector. Taking the Laplace transform of (1) and eliminating the state x, we get the $p \times m$ transfer-matrix from input u to output y

$$Y(s) = G(s)U(s) \; ; \quad G(s) = C(sI - A)^{-1}B + D \tag{2}$$

We will also make extensive use of the frequency response of the system which is obtained by letting $s = i\omega$ in $G(s)$

$$Y(i\omega) = G(i\omega)U(i\omega) \tag{3}$$

where $Y(i\omega)$ and $U(i\omega)$ are the Fourier transforms of $y(t)$ and $u(t)$, respectively.

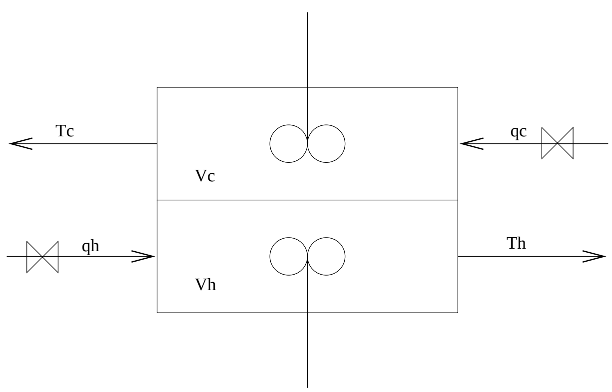

Figure 1.2: Simple heat exchanger with hot q_h and cold q_c flows.

As an example, consider the simple heat exchanger in Figure 1.2, used to transfer heat from a hot stream q_h to a cold stream q_c. Energy balances across the hot and cold side, respectively, yield the differential equations for the temperatures T_h and T_c.

$$\begin{aligned} V_h \frac{dT_h}{dt} &= \alpha_h(T_c - T_h) + q_h \beta_h (T_{hi} - T_h) \\ V_c \frac{dT_c}{dt} &= \alpha_c(T_h - T_c) + q_c \beta_c (T_{ci} - T_c) \end{aligned} \tag{4}$$

A linear time invariant model is obtained by linearizing this model about a desired steady-state. Assume $V_h = V_c = 1$, $\alpha_h = \alpha_c = 0.1$, $\beta_h = \beta_c = 0.5$, $T_{hi} = 100C, T_{ci} = 20C$,

$q_h = 0.1m^3/s$ and $q_c = 0.04m^3/s$. At steady-state the time-derivatives are zero, which yields the steady-state values $T_h^* = 80C$ and $T_c^* = 70C$. A Taylor series expansion of the right hand sides of (4) about this steady-state, keeping only the linear terms, yields

$$\begin{aligned} \dot{x} &= \begin{pmatrix} -0.15 & 0.1 \\ 0.1 & -0.12 \end{pmatrix} x(t) + \begin{pmatrix} 10 & 0 \\ 0 & -25 \end{pmatrix} u(t) + \begin{pmatrix} 0.05 & 0 \\ 0 & 0.02 \end{pmatrix} d(t) \\ y(t) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x(t) \end{aligned} \tag{5}$$

where the state vector $x = [\Delta T_h \ \Delta T_c]^T$, the input $u = [\Delta q_h \ \Delta q_c]^T$, the disturbance $d = [\Delta T_{hi} \ \Delta T_{ci}]$ and the output $y = [\Delta T_h \ \Delta T_c]^H$. Note that all variables in a linear model are deviation variables, e.g., $y_1 = \Delta T_h = T_h - T_h^*$, corresponding to the deviation from the steady-state value at which the linearized model was obtained. Taking the Laplace transform $G(s) = C(sI - A)^{-1}B$, we obtain

$$G(s) = \frac{1}{(29.5s + 1)(4.24s + 1)} \begin{pmatrix} 150(8.33s + 1) & -312.5 \\ 125 & -469(6.67s + 1) \end{pmatrix}$$

for the transfer-matrix from the input u to the output y. As an exercise you can derive the corresponding transfer-matrix $G_d(s)$ from the disturbance d to the output y. Note that, with reference to the block-diagram in Figure 1.1, the disturbance on the output is $w = G_d(s)d$.

Remark: Note that the controllers $F_y(s)$ and $F_r(s)$ we design in this course in general will be LTI models of the controllers, just like $G(s)$ and $G_d(s)$ are models of the physical systems we want to control. Thus, F_y and F_r must be *realized* in order to implement them on the real system. That is, we need to create a real system that has the dynamics corresponding to the models $F_y(s)$ or $F_r(s)$. The realization is in most cases solved by implementing the controller as an algorithm in a digital computer. Realization of the controller will not be covered in any detail in this course, but is covered e.g., in the course *EL2450 Hybrid and Embedded Control Systems.*

One may ask why we usually rely on feedback in control systems? We will in this course for the most part assume that a model is available for the plant G

$$z = G(s)u + w \tag{6}$$

Combining (6) with the control objective $z = r$, we get

$$u = G^{-1}(r - w) \quad \Rightarrow \quad z = r \tag{7}$$

which corresponds to feedforward control from the reference r and disturbance w. Thus, if we have a perfect model and we measure the disturbance w then there appears to be no need for feedback control. However, there are at least three reasons why we usually need feedback.

First, there will always be some mismatch between the true system and the model; we term this *model uncertainty.* Assume for instance that the true system $\tilde{G}$ is given by

$$\tilde{G} = G(I + \Delta_G)$$

where G is the model and Δ_G is the relative model error, i.e., the relative difference between the true plant and the model. Combining this with the feedforward control (7) we get

$$z = r + G\Delta_G G^{-1}(r - w)$$

Thus, for $\Delta_G \neq 0$ we do not get perfect following of the setpoint r. In fact, as we shall see later in the course, the second term representing the control error $e = r - z$ can become very large even for relatively small model errors Δ_G. Feedback is an efficient way of reducing the impact of model uncertainty. In Lecture 2 we will quantify the effect of feedback on uncertainty.

Second, there are usually many different disturbances acting on a system and we seldom measure all of these, if any at all. Unmeasured disturbances can be seen as uncertainty about the environment of the system we want to control and can only be dealt with using feedback control. Note that without knowledge of w and no feedback we have $z = w$, i.e., the disturbance goes directly through to the output. In Lecture 2 we will also quantify the effect of feedback on disturbance attenuation.

Finally, if the system $G(s)$ is unstable then the only way to make the system stable is by the use of feedback.

In summary, the main reasons for employing feedback, as opposed to feedforward, control are

- Model uncertainty (uncertain knowledge about system)
- Unmeasured disturbances (uncertain environment)
- Unstable systems

The main costs of feedback are partly that the controller F_y feeds measurement noise into the plant via the control input u, and partly that we potentially risk inducing instability in an otherwise stable system.

Note that we in general can combine feedforward and feedback control, and that for instance the pre-filter F_r in Figure 1.1 is a feedforward control from the setpoint r.

1.1 Controller Design

The classical control design and analysis methods, initially developed by pioneers such as Bode and Nyquist and typically dealt with in an introductory course in control, are valid for single-input-single-output (SISO) systems only. In this course we deal with multi-input-multi-output (MIMO) systems for which the classical methods are not directly applicable. However, it may be tempting to consider a MIMO system as a collection of SISO systems, and control each output with one input. We consider this case first to explain why it usually is not a viable approach.

As an example, consider the heat-exchanger above. We may consider controlling the hot temperature $y_1 = T_h$ using the hot flow $u_1 = q_h$ and the cold temperature $y_2 = T_c$ using the cold flow $u_2 = q_c$. If we for instance employ PI-controllers we then have

$$u_1 = \underbrace{K_{c1}\frac{T_{i1}s+1}{T_{i1}s}}_{C_1}(r_1 - y_1)\ ; \quad u_2 = \underbrace{K_{c2}\frac{T_{i2}s+1}{T_{i2}s}}_{C_2}(r_2 - y_2)$$

This control strategy is called *decentralized control* and is illustrated by the block diagram

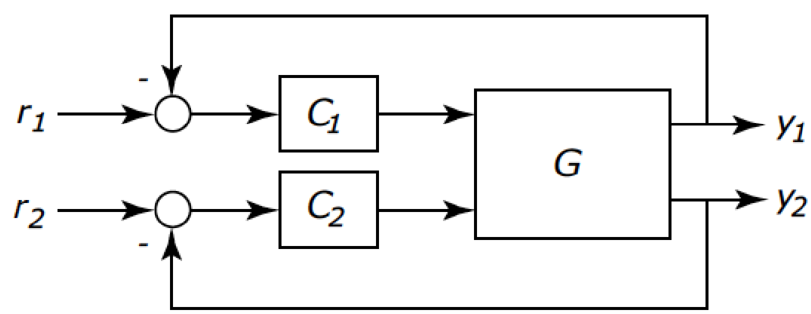

Figure 1.3: Decentralized control of a 2×2 system

in Figure 1.3. If we now consider each loop separately, we can tune these using classical control design methods. For the individual loops we then get the closed-loop transfer-functions

$$y_1 = \underbrace{\frac{G_{11}C_1}{G_{11}C_1 + 1}}_{G_{c1}} r_1\ ; \quad y_2 = \underbrace{\frac{G_{22}C_2}{G_{22}C_2 + 1}}_{G_{c2}} r_2$$

However, since both inputs affect both outputs, the two loops will interact when we close both of them and the behavior will in general not be as expected from considering the individual loops (see also example 1.1 in course book or slides from Lecture 1).

If we consider the problem on the correct multivariable (matrix) form we get

$$y = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix} (r - y) \quad \Rightarrow \quad y = \underbrace{(I + GC)^{-1}GC}_{G_c} r$$

where $y = [y_1\ y_2]^T$ and $r = [r_1\ r_2]^T$. Since G is a full matrix, so will also G_c be full and the diagonal elements of G_c will in general be very different from the individual elements G_{c1} and G_{c2}. In fact, G_c may be unstable even if the individual loops are stable. Furthermore, the off-diagonal elements of G_c will in general be non-zero, implying that changing the setpoint for one of the outputs will cause both outputs to change. Thus, in the general case it is not advisable to treat MIMO problems as a collection of SISO systems. Rather, one should approach them as multivariable systems as such, i.e., based on transfer-matrices and signal vectors. Since the classical design methods, like lead-lag design, are limited to SISO systems we will need to employ design and analysis methods that are specifically aimed at multivariable systems. One convenient approach to design controllers for multivariable systems is to formulate the control problem as an optimization problem, i.e., minimize the control error e and input usage u. For this we will need to quantify the size of signals and systems, and for this purpose we introduce below the concept of *signal and system norms.*

Note that there may be systems where a diagonal controller C, i.e., decentralized control, will work satisfactory. We will later in the course consider methods for determining when this will be the case.

1.2 Signal Norms and System Gain

In multivariable dynamic systems, all signals are vectors that furthermore are functions of time (or frequency). If we want to keep a signal, e.g., the control error e, small in some sense, we need to be able to quantify its size. A *vector norm* is a function that maps a vector into a scalar positive number, while a *signal norm* maps a function of time or frequency into a scalar positive number. Norms have to satisfy certain basic properties, but we will not go into any details here but rather define the specific norms that we will use in this course. For a more in-depth introduction to norms, we refer to a course on functional analysis.

There exists many different norms that can be used to quantify the size of a vector. In this course, we we will only use the most common vector norm, namely the Euclidian 2-norm. For a real vector $z \in \mathbb{R}^m$ the vector 2-norm is defined as

$$|z| = \sqrt{\Sigma_{i=1}^m z_i^2} = \sqrt{z^T z}$$

Note that this is the standard definition of the (Euclidian) length of a vector.

A signal norm measures the size of a time varying signal. If the signal is a vector, then a vector norm is used to convert it into a scalar time varying signal. The L_∞-norm, or peak-norm, is defined as

$$\|z\|_\infty = \sup_{t \geq 0} |z(t)|$$

where $|\cdot|$ denotes the Euclidian 2-norm (sup denotes the supremum, or least upper bound, which for most practical cases equals the maximum value). A signal is said to be bounded if the peak-norm is finite, i.e., $\|z\|_\infty < \infty$.

The L_2-norm, or energy-norm, is defined as

$$\|z\|_2 = \sqrt{\int_{-\infty}^{\infty} |z(t)|^2 dt}$$

A signal is said to be finite-energy if the energy-norm is finite, i.e., $\|z\|_2 < \infty$.

We will in this course only use the L_2-norm, or 2-norm for short.

A system $\mathcal{S}$ can be seen as a mapping of an input signal u to an output signal y. See Figure 1.4. Just like we use norms to obtain scalar measures of the size of signals, we

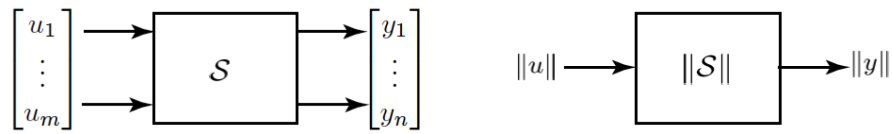

Figure 1.4: A system is a mapping of input signals to output signals. A norm $\|\cdot\|$ is function that assigns a positive real number, quantifying the size, to a time-varying or frequency dependent signal, or to the amplification of a system.

would like to have a scalar measure for the amplification of a system. For this purpose we employ system norms, also known as system gain. Consider first that we employ a specific input $u(t)$ which gives a specific output $y(t) = \mathcal{S}u(t)$. If we use the energy-norm, introduced above, to measure the size of the signals then the amplification exerted by the system is

$$\frac{\|y\|_2}{\|u\|_2} = \frac{\|\mathcal{S}u\|_2}{\|u\|_2}, \ \|u\|_2 \neq 0$$

The amplification will depend on the specific signal considered. To obtain a scalar measure of the system, we consider the maximum amplification over all possible signals

$$\|\mathcal{S}\| = \sup_{\|u\|_2 \neq 0} \frac{\|\mathcal{S}u\|_2}{\|u\|_2}$$

The maximum amplification, when both input and output are measured in the 2-norm, is called the energy-gain of the system, or simply the system gain.

The above definition of the energy-gain is valid for any system. Let us consider a stable linear time invariant SISO system with transfer-function $Y(s) = G(s)U(s)$, for which the frequency response is $Y(i\omega) = G(i\omega)U(i\omega)$ where $Y(i\omega)$ and $U(i\omega)$ are the Fourier transforms of $y(t)$ and $u(t)$, respectively. Assume that the peak amplitude of $|G(i\omega)|$ is K, i.e., $|G(i\omega)| \leq K \ \forall\omega$, and that $|G(i\omega^*)| = K$, i.e., the peak occurs at the frequency $\omega = \omega^*$. Then the energy-norm of $y(t)$ is, using Parseval's Theorem (the integral of the square of a function is equal to the integral of the square of its Fourier transform),

$$\|y\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |Y(i\omega)|^2 d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G(i\omega)|^2 |U(i\omega)|^2 d\omega \leq K^2 \|u\|_2^2$$

Note that equality holds if $u(t) = \sin(\omega^* t)$ and hence the system gain is

$$\|G\| = \sup_{\omega} |G(i\omega)| = \|G\|_\infty$$

The energy-gain of a linear time invariant system G is called the H_∞-norm and is denoted $\|G\|_\infty$. Thus, for a stable SISO system it is simply the peak value of the amplitude $|G(i\omega)|$ in the Bode plot of G.

Consider next a simple static nonlinear nonlinear system

$$\mathcal{S} \ : \ y(t) = f(u(t)) \ ; \quad |f(x)| \leq K|x|$$

and $|f(x^*)| = K|x^*|$. Then the energy-norm of the output is

$$\|y\|_2^2 = \int_{-\infty}^{\infty} |f(u(t))|^2 dt \leq \int_{-\infty}^{\infty} K^2 |u(t)|^2 dt = K^2 \|u\|_2^2$$

and hence the energy-gain is

$$\|\mathcal{S}\| = \sup_u \frac{\|y\|_2}{\|u\|_2} = K$$

Since we in this course will focus on MIMO systems, we need to extend the definition of gain to transfer-matrices. We start by considering the gain for a static linear MIMO system $y = Au$. Then, the gain is

$$\|A\| = \sup_{u \neq 0} \frac{|y|}{|u|} = \sup_{u \neq 0} \frac{|Au|}{|u|}$$

Taking the square

$$\|A\|^2 = \sup_{u \neq 0} \frac{|Au|^2}{|u|^2} = \sup_{u \neq 0} \frac{u^T A^T A u}{u^T u} = \lambda_{max}(A^T A)$$

where λ_{max} is the maximum eigenvalue of $A^T A$. Hence, the gain of a static linear system (constant matrix) A is the square root of the maximum eigenvalue of $A^T A$. The square roots of the the eigenvalues of $A^T A$ are called the singular values of A and are denoted $\sigma_i(A)$. The largest singular value, which is the gain of A, is denoted $\bar{\sigma}(A)$, i.e.,

$$\|A\| = \bar{\sigma}(A)$$

We will return to the singular values later when we discuss properties of MIMO systems in more detail.

Using the results above we can show that the gain for MIMO LTI system G is

$$\|G\| = \sup_{\omega} \bar{\sigma}(G) = \|G\|_{\infty}$$

i.e., the peak value of the maximum singular value over all frequencies.

1.3 The Small Gain Theorem

The small-gain theorem is a general and very useful result on the stability of feedback systems that we will make extensive use of in this course. Essentially, it states that if the open-loop system is stable and the loop-gain is less than one, then also the closed-loop is stable. The definition of stability we use here is that of *input-output stability*. A system is said to be input-output stable if a bounded energy input gives a bounded energy output, or if the gain

$$\|\mathcal{S}\| < \infty$$

that is, the system has finite gain. We can then formally formulate the small gain theorem

The Small Gain Theorem: Consider the feedback interconnection of two systems $\mathcal{S}_1$ and $\mathcal{S}_2$ in Figure 1.5, where $\mathcal{S}_1$ and $\mathcal{S}_2$ both are input-output stable and the loop gain

$$\|\mathcal{S}_1\|\|\mathcal{S}_2\| < 1$$

Then, the closed-loop system is input-output stable from any input r_1, r_2 to any output e_1, e_2, y_1, y_2.

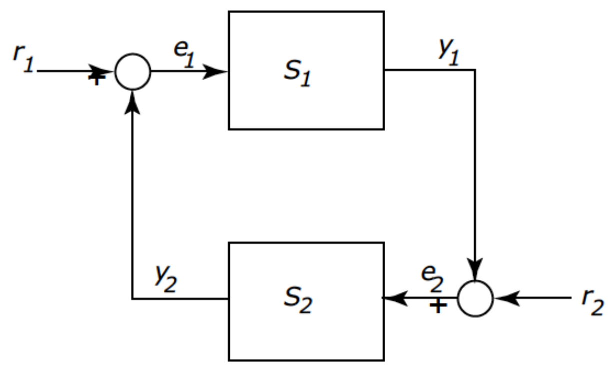

Figure 1.5: Feedback interconnection of two stable systems $\mathcal{S}_1$ and $\mathcal{S}_2$.

The proof is based on the Generalized Nyquist Theorem which we will introduce later in the course, and the proof is therefore left out for now. An informal sketch to a proof is given on the slides for Lecture 1.

Note that this is a sufficient condition for stability only, and by no means necessary. Consider for instance a linear negative SISO feedback loop with stable loop transfer-function $L(s)$. Then, from the Bode stability criterion the closed-loop is stable if the amplification $|L(i\omega)| < 1$ at the frequency where the phase lag is $\arg L(i\omega) = -\pi$. Thus, if we require $|L(i\omega)| < 1$ at all frequencies, as in the small gain theorem, then clearly we have closed-loop stability. But, this is then potentially highly conservative since we according to the Bode criterion can have $|L(i\omega)| >> 1$ at frequencies where the phase lag is not $-\pi$. Despite this conservativeness, we shall later see that the Small Gain Theorem is very useful, in particular when analyzing robustness, i.e., stability in the presence of uncertainty.