ver 1.1, 190326 Elling W. Jacobsen, KTH

1 EL2520 Lecture notes 4: Fundamental Limitations and Conflicts

In many control courses the focus is mainly on control design, and analysis of the resulting closed-loop system. With this approach it is easy to get led to believe that the control performance relies only on the ability of the control engineer to design a sufficiently good controller. Also, one can get the impression that if one cannot meet the desired performance with one control design, then one should simply try with a more advanced controller.

The fact is that there always are hard limits to what can be achieved with feedback for a given system, and these limits are set by the system itself. Thus, the only way to overcome the limits, given that they prevent acceptable control performance, is a redesign of the system itself. This underlines that the traditional sequential approach to system and control design, i.e., first design the system then the controller to achieve desired dynamic behavior, is not optimal. Rather, the two activities should be integrated to ensure that the desired performance can be achieved.

In this lecture we derive quantitative limitations on the achievable control performance for a given system. We will mainly consider performance in terms of the sensitivity function S and the complementary sensitivity function T. As before, we specify the performance using weights W_S and W_T, respectively, so that the design requirements are

$$\|W_S S\|_\infty \leq 1 \quad \Rightarrow \quad |S(i\omega)| < |W_S^{-1}(i\omega)| \; \forall\omega$$
$$\|W_T T\|_\infty \leq 1 \quad \Rightarrow \quad |T(i\omega)| < |W_T^{-1}(i\omega)| \; \forall\omega$$

A standard choice of weight for the sensitivity function S is

$$W_S = \frac{1}{M_s} + \frac{\omega_{BS}}{s} \tag{1}$$

corresponding to infinite weight at $\omega = 0$ (enforcing $S(0) = 0$), weight $|W_S(i\omega_{BS})| \approx 1$ (enforcing minimum bandwidth for S approx. ω_{BS}) and, finally, maximum peak of $|S|$ equal to M_S at frequencies above ω_{BS}. Similarly, a standard weight for T is

$$W_T = \frac{1}{M_T} + \frac{s}{\omega_{BT}} \tag{2}$$

corresponding to allowing maximum peak M_T, minimum bandwidth ω_{BT} and enforcing $T = 0$ at infinite frequency.

1.1 Sensitivity Trade Off

With loop-gain $L = GF_y$ we have

$$S = \frac{1}{1+L} \; ; \quad T = \frac{L}{1+L}$$

and hence

$$S(i\omega) + T(i\omega) = 1 \quad \forall \omega$$

This implies that either $|S(i\omega)| > 0.5$ or $|T(i\omega)| > 0.5$ at any frequency, that is, we can not deal effectively with both disturbances and noise at the same frequency. In terms of the weights, a consequence is that we can not choose $|W_S| > 2$ and $|W_T| > 2$ at the same frequency.

Also note that since $|S(i\omega)+T(i\omega)| = 1$, the distance between $S(i\omega)$ and $-T(i\omega)$ is always 1 and hence

$$|S(i\omega)| >> 1 \quad \Leftrightarrow \quad |T(i\omega)| >> 1$$

at any given frequency ω. Thus, a large peak in $|S|$ implies a large peak in $|T|$, and vice versa.

1.2 The Bode Sensitivity Integral

In Lecture 2 we showed by simple arguments, based on the Nyquist plot of the loop gain, that the sensitivity function of any closed-loop systems must exceed $|S| > 1$ at some frequency. A more powerful result is due to Bode:

Bode Integral Theorem: Suppose that the loop-gain $L(s) = GF_y(s)$ has relative degree (pole excess) ≥ 2, and that $L(s)$ has N_p poles in the RHP located at $s = p_i$. Then the sensitivity function must satisfy

$$\int_0^\infty \log|S(i\omega)|d\omega = \pi \Sigma_{i=1}^{N_p} Re(p_i)$$

The proof is based on Cauchy integral theorem (see course on complex analysis).

Thus, if we plot $\log|S|$ vs ω, then the area for which $|S| > 1$ ($\log|S| > 0$) must at least equal the area for which $|S| < 1$ ($\log|S| < 0$). If the system is open-loop stable ($N_p = 0$) then the two areas must exactly match. Thus, if we push down the sensivitiy in one frequency range we must pay for this by increasing the sensitivity in another frequency range. This is commonly known as the *waterbed effect* in control. Thus, we are enforced to make a trade-off between the performance in different frequency ranges.

1.3 Interpolation Constraints from RHP Poles and Zeros

Consider again the loop-gain

$$L(s) = G(s)F_y(s)$$

If $G(s)$ has a zero at $s = z$ in the complex RHP, then $G(z)$=0 and $L(z) = 0$. The last equality follows from the fact that we are not allowed to cancel RHP zeros in G by corresponding RHP poles in the controller $F_y(s)$ since then we do not get internal stability (see Lecture 2). The implications for the sensitivity functions are

$$S(z) = \frac{1}{1+L(z)} = 1 \; ; \quad T(z) = \frac{L(z)}{1+L(z)} = 0 \tag{3}$$

Similarly, if $G(s)$ has a pole at $s = p$ in the complex RHP, then $G(p) = \infty$ and $L(p) = \infty$, and

$$S(p) = \frac{1}{1 + L(p)} = 0 \; ; \quad T(p) = \frac{L(p)}{1 + L(p)} = 1 \tag{4}$$

These are interpolation constraints on $S(s)$ and $T(s)$ that any stabilizing controller must satisfy.

To see what the interpolation constraints above implies for the bounds on $|S(i\omega)|$ and $|T(i\omega)|$ we need the following result, known as the maximum modulus theorem in complex analysis

Maximum Modulus Theorem: Suppose that Ω is a region in the complex plane and that F is an analytic function on Ω and, furthermore, that F is not equal to a constant. Then $|F|$ has its maximum value at the boundary of Ω.

For a proof, see a course book on complex analysis.

Since S and T are stable transfer-functions, they have no singularities (poles) in the RHP and are hence analytic in complex RHP for which the boundary is the $i\omega$-axis. A trivial consequence of (3)-(4) and the maximum modulus theorem is then

$$\|S\|_\infty > |S(z)| = 1 \; ; \quad \|T\|_\infty > |T(p)| = 1 \tag{5}$$

The bounds in (5) are not too useful as such since all they say is that we must have a peak in $|S|$ and $|T|$ exceeding 1, which we already knew applies to essentially any system, also those with no RHP poles and zeros (see Lecture 2). Too obtain more informative bounds, we include weights on S and T.

Consider first the weighted sensitivity for which the interpolation constraint and maximum modulus theorem implies

$$\|W_S S\|_\infty > |W_S(z)S(z)| = |W_S(z)|$$

Thus, in order to achieve $\|W_S S\|_\infty \leq 1$ the weight must fulfill the constraint

$$|W_S(z)| \leq 1$$

To get some insight into what this constraint implies, consider the weight (1)

$$W_S(s) = \frac{1}{M_S} + \frac{\omega_{BS}}{s}$$

where ω_{BS} is the required bandwidth, i.e., the frequency where $|S|$ becomes (approximately) larger than 1, and M_S the maximum allowed peak of $|S|$. Then

$$|W_S(z)| \leq 1 \quad \Rightarrow \quad \frac{1}{M_S} + \frac{\omega_{BS}}{z} \leq 1$$

and from this we derive

$$\omega_{BS} \leq (1 - M_S^{-1})z$$

Thus, if we allow a peak $M_S = \infty$ then the maximum bandwitdh $\omega_{BS} = z$. A more reasonable maximum peak is $M_S = 2$ for which we get the bandwidth limitation

$$\omega_{BS} \leq \frac{z}{2} \tag{6}$$

Thus, a RHP zero places a hard limitation on the achieveable bandwidth of a control system and the limitation is worse the closer the zero is to the imaginary axis.

Example 1: Consider the plant

$$G(s) = \frac{s-1}{(s+1)(s+2)}$$

Since the plant has a RHP zero at $z = 1$, we can not achieve a bandwidth for the sensitivity larger than 1 rad/s, or we if allow a maximum peak $M_S = 2$, not larger than 0.5 rad/s. Thus, we can not attenuate disturbances at higher frequencies.

A RHP zero is known as *non-minimum phase* since there exist transfer-functions with the same amplitude $|G(i\omega)|$ but with less negative phase. For instance, the plant $G = 1/(s+2)$ has the same amplitude curve as the plant in Example 1, but more positive phase. More negative phase implies that a system is harder to control. Another non-minimum phase phenomena is a time-delay of θ, with transfer-function $e^{-\theta s}$. By employing a 1st order Padé approximation of a delay we get

$$e^{-\theta s} \approx \frac{1 - \frac{\theta}{2}s}{1 + \frac{\theta}{2}s}$$

and hence a time-delay of θ can be seen as a RHP zero at $z = 2/\theta$. Then, the bounds derived above for a RHP zero yields, with $M_S = 2$,

$$\omega_{BS} \leq \frac{1}{\theta}$$

Consider next the interpolation constraint $T(p) = 1$ at a RHP pole p of $G(s)$. With the introduction of the weight W_T, the maximum modulus theorem yields

$$\|W_T T\|_\infty > |W_T T(p)| = |W_T(p)|$$

and hence we require

$$|W_T(p)| \leq 1$$

to achieve $\|W_T T\|_\infty \leq 1$. With the specific weight (2) we get

$$\frac{p}{\omega_{BT}} + \frac{1}{M_T} \leq 1$$

and with $M_T = 2$ we get the bandwidth limitation

$$\omega_{BT} \geq 2p \tag{7}$$

Note that this is a lower bound on the bandwidth of T and is as expected since we need a minimum bandwidth to stabilize an unstable system.

Given the fact that a RHP zero imposes an upper bound on the bandwidth of S and a RHP pole imposes a lower bound on the bandwidth of T, and these bandwidths must be close since $S + T = 1$, we would expect difficulties if a plant has both RHP zeros and RHP poles and these are close. To quantify this, recall that $S(p) = 0$ at a RHP pole p. Factorizing the sensitivity function

$$S = S_{mp} \underbrace{\frac{s-p}{s+p}}_{S_{ap}}$$

where S_{mp} has no poles or zeros in the RHP and S_{ap} is all-pass, i.e., has amplitude 1 $\forall\omega$. If $G(s)$ also has a RHP zero at $s = z$, then from the constraint $S(z) = 1$ we get

$$S_{mp}(z) = S_{ap}^{-1}(z) = \frac{z+p}{z-p}$$

Hence

$$\|W_S S\|_\infty = \|W_S S_{mp}\|_\infty \geq |W_S(z) S_{mp}(z)| = |W_S(z)\frac{z+p}{z-p}|$$

For instance, with weight $W_S = 1$ we obtain

$$\|S\|_\infty \geq \frac{|z+p|}{|z-p|}$$

Thus, with poles and zeros close in the RHP in $G(s)$, large peaks in $|S|$ is unavoidable.

For the complementary sensitivity T we have $T(z) = 0$ and factorizing as for S above we get

$$\|W_T T\|_\infty \geq |W_T(p)\frac{p+z}{p-z}|$$

and with $W_T = 1$ we get

$$\|T\|_\infty \geq \frac{|z+p|}{|z-p|}$$

and large peaks in $|T|$ must also exist with z and p close, even with the best possible controller.

Example 2: Consider the inverted pendulum placed on a cart in Figure 1. The aim is to stabilize the pendulum in an upright position using the movement of the wagon by the drag force F. Consider the position x of the wagon as the output, then the transfer-function from the input F is

$$X(s) = \frac{ls^2 - g}{s^2(Mls^2 - (M+m)g)} F(s)$$

The system has both a RHP zero and a RHP pole

$$z = \sqrt{\frac{g}{l}} \; ; \quad p = z\sqrt{1 + m/M}$$

For a pendulum with length $l = 1$ and mass of the pendulum equal to the mass of the wagon, i.e., $m = M$, we get $z = \sqrt{10}$ and $p = \sqrt{20}$ and

$$\|S\|_\infty \geq 5.8 \; ; \quad \|T\|_\infty \geq 5.8$$

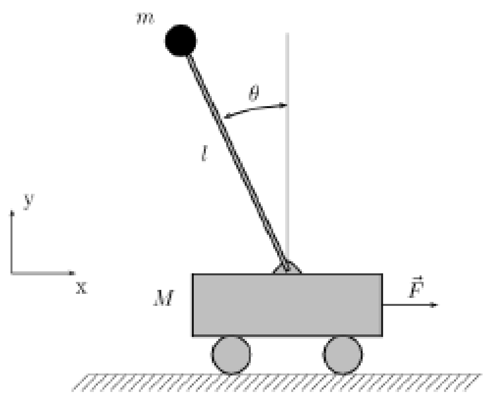

Figure 1.1: Inverted pendulum on a cart

If we reduce the mass of the pendulum to one tenth, i.e., $m = 0.1M$, we get $z = \sqrt{10}$ and $p = \sqrt{11}$ and

$$\|S\|_\infty \geq 42 \; ; \quad \|T\|_\infty \geq 42$$

Note that the rocket stabilization problem that puzzled scientists in the late 1950s and early 1960s corresponds to the cart pendulum problem since the air under the rocket becomes fluidized. From the result above one can understand why the problem was challenging. This is a good example of a system where it is crucial to consider integrated design of system and controller to ensure satisfactory dynamic performance.

In summary, we have shown that there exist several trade-offs and inherent limitations in control, and these are fundamental in the sense that no controller can overcome them. Thus, if a limitation is in conflict with the desired dynamic performance of a system, then the only viable solution is to redesign the system to reduce the limitations. Note that all results presented here are valid for SISO systems only. In Lecture 6 we will extend the results to the MIMO case.