Chapter 7

Feedback

We need feedback
to know how well we have done
and to improve our future behavior.
But should it be positive or negative feedback?

7.1 Introduction and Motivation

The behavior of two or more systems connected in series (or cascade) is relatively easy to predict. The first system input is a free signal. Its output is the input to the next system, which in turn responds to this input, and so on, till we reach the end of the cascade. The behavior of a cascade is well-defined, as long as each sub-system in the cascade is well-defined[1]. At each stage the signal is affected by the system in that stage. Despite its simplicity, we know that a cascade does not always deliver as expected, as anyone who has ever played the game of *Chinese whispers*[2] knows all too well.

When systems are interconnected in a loop things are different. Even with two systems in a loop, where do we start? When do we stop chasing our own tail? Predicting what behavior is going to exist is no longer a matter of simply understanding the subsystems in their own right. The interconnection is essential in defining the behavior. Moreover, where typically cascade connections are always well-defined, feedback connections are not. A well-known example of a feedback loop that is not well-posed is the acoustic feedback that occurs with having a microphone listening to the output of a speaker. Depending on the amplifier in the loop, the result can be most unpleasant. The shower example, illustrated at the start of this chapter, is another example of how feedback has to be treated with care.

Just consider two unitary systems interconnected in a positive feedback loop with external additive input $r = 1$. Doing some simple calculation you will come to the absurdity that $1 = 0$!!!, indicating that the feedback system is not well-posed. A little more generally, as discussed in Sect. 5.6.2, the basic operator in a simple feedback loop consisting of two systems G_1 and G_2, is given by Eq. 5.13, rewritten here

$$G = G_1(1 + G_1G_2)^{-1} \tag{7.1}$$

The issue of well-posedness is concerned with the existence of $(1 + G_1G_2)^{-1}$. In our previous example this expression is $(1 + (-1)(1))^{-1}$ for which there is no valid interpretation.

---

[1] Tacitly, we are assuming that the outputs of the previous system in the cascade are acceptable inputs to the next system, but presumably nobody would build or even conceive a cascade unless this was the case.

[2] Whispering a small piece of information into a neighbor's ear, who then transmits the message to the next in line and so on. At the end of the line, the received message is compared with the original. Typically the final message bears little resemblance to the initial message!

In our discussions we are going to make a leap of faith, assuming that engineers and nature know how to build systems. We are going to simply assume that the feedback loops are well-posed.

Much of the power of feedback loops has to do with the $^{-1}$ (denoting the inverse) in the above expression. A feedback loop creates some kind of an inverse or a division of an operator. Just like we cannot divide by zero (the previous example told us as much), inverses of operators are tricky, but being able to divide or take an inverse by simply interconnecting some elements is pretty neat, and very powerful.

As observed before, feedback is based on the observation of an output to modify the input to the same system. For this reason, systems incorporating feedback are also known as *reactive* systems. From a control point of view, this is a drawback because the control system in order to react must first detect the output, or error condition, before it can take action. On the other hand, very little knowledge about the system behavior may be required to react appropriately. Moreover, feedback reacts to changes in the loop, sometimes this is referred to as an adaptive response. In contrast, in open-loop (feedforward or cascade) for the final output to do what we want it to do, the models of the systems in the cascade must be precisely known, so as to select the right input signal. Roughly speaking the desired input into the cascade must equal the output processed by the inverse of the cascade, not an easy task in general. Moreover, the input to the cascade cannot react to changes in any element of the cascade.

Feedback plays a key role in systems dynamics. It is ubiquitous in both the engineered and the natural world. In this chapter we consider some of the powerful properties feedback offers. We start with a number of examples to illustrate the pervasiveness of feedback. Next we explore the power of the feedback a little more.

7.2 Internal Feedback

In Chap. 5, we learned that natural and engineered systems alike are mostly composed of rather elementary systems such as gains, integrators (accumulators) and (transport) delays, amongst many. It is precisely the existence of feedback that bestows richness on the dynamic behavior of interconnected systems. Let us revisit some of the feedback loops we already encountered.

Filling a tank. In Chap. 2, a kitchen sink (Fig. 2.7) is illustrated. The water level is stabilized because there is feedback: a constant inlet flow increases the amount of stored water until it reaches a level that produces an outlet flow equal to the input flow. We often identify this situation with a negative feedback loop: the higher the water level, the greater the outflow, that reduces this water level. See also Fig. 7.1.

Exothermal reaction. Some chemical reactions become more active as the environmental temperature increases. Some of these chemical reactions are also exothermic, that is, they release heat. Such a reaction inside a thermally well-insulated vessel will be explosive as the temperature will increase as a consequence of the reaction, which in turns accelerates the reaction, which in turn creates more heat, and so on. In this case, a positive feedback determines the dynamic evolution of the system, which in this case will lead to an explosion of some kind as to destroy the vessel, and terminate the positive feedback loop.

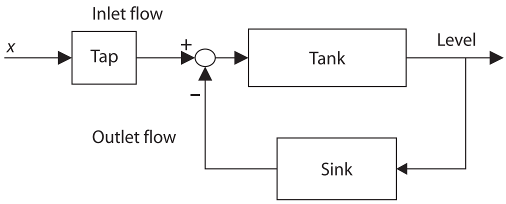

Fig. 7.1. Simple feedback loop

Electric circuit. Let us consider a simple loop circuit with a resistor, a capacitor and a battery. The voltage across the resistance is the difference between the battery voltage and that of the capacitor. As soon as a current flows through the circuit, the capacitor accumulates the extra charge and its voltage increases. Once the capacitor voltage equals the battery voltage the current ceases and the voltage stabilizes.

Motor. A voltage applied to an electric motor driving a fan generates a torque on its axis. This torque will accelerate the shaft, increasing the angular speed. A counter acting, reactive, friction torque appears as well as a mechanical load (from the wind flowing through the fan) that increases with speed. An equilibrium will be reached once the motor torque equals the mechanical counter torque.

Ecology. Remember the wolf and rabbits example in Sect. 1.4.4. On the one hand there is a positive feedback: exponential growth of the rabbit population without the predators, and on the other hand there is a negative feedback as the rabbits are eaten by the wolves. The interesting thing here is that these populations do not settle to an equilibrium, but rather evolve oscillatory, as in a limit cycle.

Internal Feedback

The dynamic behavior of a system is determined by both the dynamic behavior of the subsystems (gains, accumulators, delays and so on) as well the interconnection of these.

Feedback is intrinsic to most physical systems.

Compared to a cascade or series connection, a feedback loop made of the same components will naturally display a richer behavior.

Feedback is used to **shape** (emphasize, modify) dynamic behavior.

Using some previous examples we illustrate how feedback modifies behavior:

Thermostat. The temperature in a room is stable when heat lost balances against heat gained. A thermostat will switch a heater on when the room is too cold and when the room reaches the right temperature it switches the heater off. This rather trivial on/off, negative feedback, controller based on some measurement of the room temperature is able to regulate the room temperature, as long as the heater is sufficiently powerful to reach the set temperature for the given room dynamics as determined by its size, the outside temperature, and the insulation (windows open or not).

Motion control. In playing any ball sport, our eyes provide positional feedback and assist us in directing our motion to intercept the ball, and/or provide it with the desired motion. Similarly in driving a car we use preview information from the road ahead of us to steer the car, but we rely on our reflexes (fast control loop) to avoid the dog that runs across the street.

From these examples we realize that we do not need a lot of information, nor a precise model of the plant behavior, to implement feedback (control). There is even model-free control based on feedback (like the thermostat). In many *simple* systems, negative feedback usually has a stabilizing effect (as in the tank filling, or the rotating motor) whereas positive feedback tends to make a system unstable (as in the exothermic reaction or the rabbit's colony from Chap. 1, or indeed the acoustic feedback).

In general though, to predict the dynamic behavior of more complex systems in a loop or with multiple loops is not straightforward. Feedback may create instabilities, but it also enables performance and behavior that is nearly impossible to arrive at otherwise.

Also, positive feedback is not universally bad. For example, oscillators, like clocks, are important and in biology as in the engineered world, positive feedback is exploited to create oscillators. In social networks, positive feedback encourages and can engender desired behavior in those that receive it (we all love praise).

7.3 Feedback and Model Uncertainties

The operational amplifier. An operational amplifier is an electronic circuit with a very large gain A (say $A > 10^5$) between its input and output voltage. The normal output voltage is measured in Volts, say less than 10 V and hence the input voltage will be less than 0.1 mV (10 V divided by a gain of more than 10^5).

When building circuits with operational amplifiers, we can rely on the fact that the gain for any one amplifier is large, but we cannot rely on the precise value of this gain. Indeed the variation of the gain across operational amplifiers even from the same manufacturing batch is also very large. A gain variation of a factor of 10 or more is not uncommon. Also, over the life of the circuit, the gain will change with time and the temperature of the circuit. The redeeming feature is that the gain stays large no matter what (say somewhere between 10^4 and 10^7). It also follows that that under normal behavior, the input voltage to the operational amplifier will always be negligible (less than 1 mV).

If an operational amplifier is connected in a negative feedback configuration, that is, the output voltage is looped back to the terminal labeled with a $-$, as shown in Fig. 7.2a, the gain in the new circuit, between the output V_o and the input V_i is:

$$\frac{V_\mathrm{o}}{V_\mathrm{i}} = -\frac{R_\mathrm{o}}{R_\mathrm{i}} \frac{1}{1 + \frac{1}{A}\left(1 + \frac{R_\mathrm{o}}{R_\mathrm{i}}\right)} \tag{7.2}$$

This relationship is represented in the equivalent block diagram Fig. 7.2b. Because the amplifier's gain A is very large no matter what, the input-output relationship is well approximated by

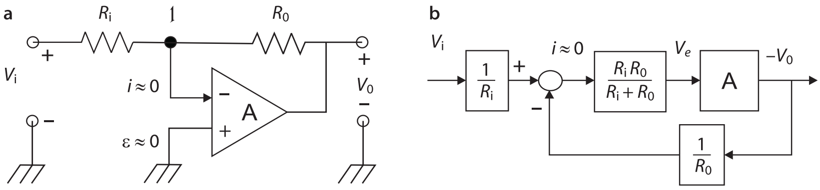

Fig. 7.2. An operational amplifier

$$\frac{V_o}{V_i} \simeq -\frac{R_o}{R_i}$$

Observe that the actual gain of the amplifier does not matter, nor does the rest of the circuit, where V_o or V_i are connected into! (Amplifiers simplify circuit analysis and synthesis considerably!) The relationship between these two voltages is solely based on the resistors around the operational amplifier. When these resistors can be defined very accurately, so is the relationship between the input and output voltage. This is an important simplification, and all due to large negative feedback.

The Operational Amplifier

For those who know some circuit theory. With Ohm's law (the current through a resistor is proportional to the voltage drop across a resistor), and Kirchoff's laws, (that the sum of voltages in a loop is zero and the sum of currents at a junction is zero), you can readily derive the input-output relationship for the operational amplifier circuit. From Fig. 7.2a, at the summing junction 1, using Kirchoff's current law, knowing that the input current to the amplifier is zero we have

$$\frac{V_i + \epsilon}{R_i} = \frac{-\epsilon - V_o}{R_o}$$

The amplifier's equation is simply $V_o = A\epsilon$. Replacing ϵ by V_o/A in the previous equation, yields after some algebra, Eq. 7.2.

As shown in Fig. 7.2, whatever is in the direct path of the loop is not very important, as long as the gain is large enough. The whole system behaves as having a net gain of $-(R_o)/(R_i)$.

This result can be extended to any dynamic feedback system. Consider a block diagram from Fig. 6.11. Under the assumption that both the disturbance and the noise signals are zero, the following relationship holds:

$$y = \left(F\frac{GK}{1+GKH}\right)r \tag{7.3}$$

Thus, as far as the gain of the operator GK is much larger than unity, for the signals of interest, the input/output relationship is only determined by F and H (Eq. 7.4). Designing both elements accurately will deliver the desired response, independent of the plant or the feedback controller:

$$\|GK\| \gg 1 \Rightarrow y \approx FH^{-1}r \tag{7.4}$$

It may look too good to be true. In general it is, although it is almost true. Indeed we are not able to make $GK \gg 1$ and also achieve stability or well-posedness for all possible reference signals r; but it is possible to have GK large where it matters[3], when r belongs to some class of signals, and it is then possible to get the above relationship holding where r matters.

The same concept can be useful when dealing with more complex system and operators (nonlinear, stochastic, multivariable and so on).

By way of illustration, suppose that r is a constant signal, $F = H = 1$ and there is neither disturbance nor noise. In this case having in the feedback controller K a pure integrator (which is retained in GK, which is the case if the plant does not differentiate its input) would guarantee that $y = r$ whenever steady state is reached. This is true assuming a well-posed feedback. Indeed in steady state y is a constant. From Fig. 6.11 the input to the feedback controller, K is then $r - y$ also a constant. Because the feedback controller contains an integrator, this constant must be zero. If it were not so, then this constant input would be integrated and this would yield an unbounded signal inside the loop, which contradicts the steady state assumption.

Compare this reasoning with the discussion of the toilet example. The cistern is the integrator in this case, and hence by the same argument, the cistern will always be filled to the right level (regardless of the other elements in the toilet).

For this reason, feedback controllers typically contain an integrator, as the integrator enables tracking of constant reference signals without an error. This is also called regulation. Having the plant output settled at a given reference level is an important task in control design, if not the most important one. As a consequence *integral action* is included in most industrial feedback controllers. In Sect. 9.3.3 this concept is elaborated on in the so-called PID controllers[4], which are literally everywhere in the process and manufacturing industry.

Clearly feedback can provide a good (although not always a perfect) solution for tracking, provided the loop gain GK is large and $FH^{-1} \approx 1$ where it matters for the signal to be tracked. These are weak requirements compared to what an open loop solution would have to do: $F \approx G^{-1}$. The feedback solution does not need a lot of information about the plant G in order to achieve good tracking (where r is important and G is small, K should be large so as to make GK large).

7.4 System Stabilization and Regulation

As already discussed, one of the main features of feedback is the potential to stabilize an unstable system. This also means that feedback can destabilize. In the example discussed in Sect. 6.4.3, α was the feedback parameter to be tuned and we realized that the system behavior could be adjusted (stable, unstable, oscillatory, damped). The system remains unstable if α is outside the interval $\{0, -2\}$. There is only a small window of opportunity for stability!

In a more physical setting, in a typical servo motor application, where the shaft must be positioned at a particular angle (think of the radio antennae problem with a

[3] For those who may see these relationships as representations in the frequency domain, GK must be large where the spectral content of r matters.

[4] PID, proportional, integral and derivative action.

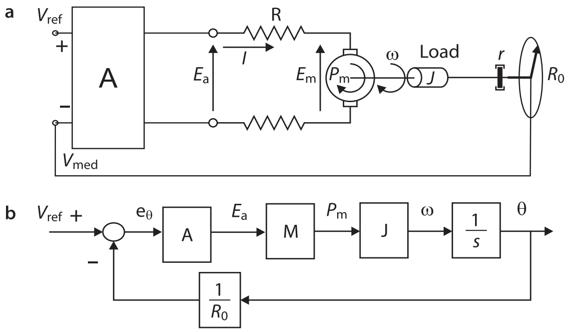

Fig. 7.3. A DC motor-based servo

fixed look direction) the variables of the system are as depicted in Fig. 7.3a. There is an electronic power amplifier/rectifier, a mechanical inertia, friction on the axis, as well as an integrator which links the motor speed to the shaft's angular position. The operator in the forward path from the drive voltage E_a to the position ϑ therefore contains an integrator. Any desired fixed angle position is clearly open-loop unstable, due to the integrator. If a voltage is applied, the motor will run, and the shaft will simply rotate. We can say that the system is unstable.

Let us introduce a negative feedback in such a way that the drive voltage is generated by, for instance, an amplifier as the one shown in Fig. 7.2a, with two inputs. The positive terminal has a voltage proportional to the desired position and the negative (feedback terminal) receives a voltage proportional to the current shaft's position. This is also shown in the block diagram Fig. 7.3b. As long as there is an error or a difference between these two inputs, a voltage will be generated by the amplifier and the motor will start to run, hopefully in the right direction so as to reduce the error (negative feedback). When the error vanishes, no voltage will be generated and the motor stops.

There will be problems if the amplifier gain is too high, because of the motor's inertia. As long as the motor receives a drive voltage it generates torque and rotates. With a large gain, even if the position error is small, this voltage is large, and the motor will have too much inertia to immediately stop or reverse direction even when the shaft went past the reference position. The drive voltage switches polarity, and eventually the motor starts turning backwards. The whole cycle repeats itself, producing unpleasant oscillations (vibrations). Feedback needs care.

As may be observed, in order to implement feedback, extra system infrastructure is needed: the components required to build the feedback path. In this servo motor example, the sensor could be just a simple potentiometer, transforming the axis position into a voltage. In other applications, the lack of appropriate measurement devices may prevent feedback from being considered.

This servo motor example, though rather different in its physical incarnation has all the features of the toilet cistern's behavior. The main features are the integrator (the motor or the cistern) and the negative feedback loop with a gain element (amplifier or float). Together these lead to a constant steady state equal to the reference (reference voltage or float position).

7.4.1 ISS and Feedback Systems

The concept of input-to-state stability considers stability of systems with inputs. This enables its use in system interconnections, and in particular feedback.

When considering a general feedback configuration of ISS systems, as depicted in Fig. 7.4 the overall system will be also ISS provided the total gain in the loop is sufficiently small. This is the celebrated *small gain* stability result.

In this figure, the external inputs are u_1, u_2, and the system states (outputs) are y_1, y_2 respectively. The internal signals are e_1, e_2. The meaning of the signals varies, but typically

- u_1 are external disturbances from the environment acting on the system of interest;
- u_2 are signals expressing control references and perhaps measurement errors;
- y_1 the state (output) of the system under control;
- y_2 the state (output) of the compensator, or control subsystem, our design freedom;
- e_1 the control input into the system under control;
- e_2 the measured feedback, derived from the system under control.

Small Gain Theorem

An important stability result associated with Fig. 7.4 goes as follows.

Assume that the ISS gain from input e_1 to output e_2 through system S_1 is given by g_1. This gain calculation is performed independent of the feedback loop structure, i.e. the input e_1 is assumed to be completely unrestricted. In the simplest case, say for linear systems, the gain tells us that the size of e_2 will be less than g_1 (a positive scalar) times the size of e_1.

Similarly assume that the ISS gain from input e_2 to output e_1 through system S_2 is g_2. Again, this gain function is determined independent of the feedback loop, g_2 is simply a property of system S_2.

Consider now the gain of the cascade S_1 after S_2 of the systems in the feedback loop. The gain of this cascade is the composition of the gain functions g_1 and g_2 and in the simplest case this gain is the product of the individual gains $g = g_1 g_2$. This gain is also called the *loop-gain*.

Conclusion. Input-to-state stability of the feedback loop, that is from the external inputs (u_1, u_2) to the external outputs/states (y_1, y_2) and also to the signals in the loop (e_1, e_2) requires that the loop-gain g is strictly less than unity.

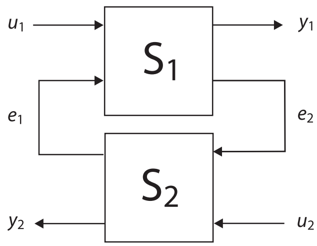

Fig. 7.4. A generic feedback loop

Gain Margin

The concepts behind the small gain theorem allow one to quantify a measure of stability for a system, or a degree of stability, denoted as the **gain margin**.

Consider in Fig. 7.4 a system G_1 (of interest) in a feedback loop, with a pure (scalar) gain system $G_2 = k$.

In general, as the gain k is increased, the feedback loop may become unstable. If the largest k such that the feedback loop is stable is larger than 1, then this is identified as the gain margin of system G_1. (1 plays a particular role, because traditionally it is assumed that the unity feedback loop is stable.)

This is perhaps counter intuitive in view of the small gain theorem, but systems may have an infinite gain margin.

More generally, the so-called conditionally stable systems are only stable for a given range of feedback gains k, say $k \in (k_{min}, k_{max})$. If zero does not belong to this interval (zero = no feedback) then such systems are open-loop unstable.

The small gain result is very useful, but also very conservative. Indeed it may well be that our estimates of the gain functions are not very tight, and these estimates may fail to meet the small-gain stability condition and yet the feedback loop may be stable and very well-behaved. Even if our gain estimates are tight, we should not expect that the small-gain condition captures all possible stable feedback loops. After all the gain functions only capture the effect of the size of a signal, and there is of course much more to a signal than just its size.

7.4.2 Linear Feedback Systems

In the case of linear systems, the question of stability and input-to-state stability can be settled using algebra for which there are efficient computational tools.

Let us reconsider Fig. 5.8 which is a special case of Fig. 7.4 for linear systems. It will become clear that feedback easily alters the stability properties and that the feedback loop can be stable and yet both systems in the loop unstable. Let the forward operator from u_1 to y_1 be $G_1(z) = N_1(z)/D_1(z)$, and the feedback operator from u_2 to y_2 be $G_2(z) = N_2(z)/D_2(z)$ (with negative feedback, as defined in Sect. 5.6.2). Using block diagram calculus, it is easy to see that the transfer function from u to $y = y_1$ is

$$G(z) = \frac{G_1(z)}{1 + G_1(z)G_2(z)} = \frac{N_1(z)D_2(z)}{N_1(z)N_2(z) + D_1(z)D_2(z)} \tag{7.5}$$

The roots of the denominator, $N_1(z)N_2(z) + D_1(z)D_2(z) = 0$ are clearly different from those of the components D_1 and D_2. That is, stability of the subsystems G_1 and G_2 does not imply stability of the feedback system. Neither does stability of the feedback system imply stability properties for the subsystems[5].

By way of example consider G_1 as a pure integrator in continuous time, $G_1(s) = 1/s$ with a feedback system which is a pure scalar gain $G_2 = g_2/1$. According to the above calculation, the closed loop stability is determined by $g_2 + s$, so that the feedback system is stable if $g_2 > 0$ (see also the discussion around Eq. 6.7). The open loop system G_1 is

[5] This will be a key point in designing feedback controlled systems, as discussed in the next chapters.

unstable. The integrator's gain margin is infinite. Applying a negative gain in G_2 has the effect of applying positive feedback, which indeed destabilizes the loop.

Similarly, in discrete time, with $G_1 = 1/(z-1)$, an integrator in discrete time and a feedback system $G_2 = g_2$ a pure gain. The feedback loop is stable provided the roots of $g_2 + (z-1) = 0$ are less than one in magnitude. The only root is $1 - g_2$, which is less than one in magnitude for $g_2 \in (0,2)$. Again, the open loop system is not stable. The gain margin is finite, it equals 2. A negative gain as well as a gain larger than 2 will lead to instability in this case (remind the example in page 161).

Bode plots and Nyquist plots examine the behavior of the loop gain $G_1 G_2$ to identify the boundary of stability and this explains to some extent the popularity of the frequency domain methods, as before computers appeared, graphical methods provided powerful design methods. The key in this frequency domain approach is to ensure that at no time the loop gain $G_1 G_2$ can have a magnitude larger than one and a phase shift of 180 degrees.

Using the frequency domain ideas, it can be seen that the instability for one or two integrators described in Sect. 5.7 require positive feedback.

This was obvious in the single integrator loop, Fig. 5.14 where we explicitly used positive feedback to obtain an unbounded response (as explained just above again). Using frequency domain ideas, considering sinusoidal signals as input, it is clear that the integrator produces a sinusoidal output, of the same frequency, but with a phase shift equal to $-\pi/2$,[6] regardless of the frequency. It then follows that the loop gain can be arbitrarily large, and no instability will develop.

Positive feedback can however occur in a loop, even when we think we are applying negative feedback. As previously discussed, the elements in the loop will determine a loop gain and phase shift. Assume that for some frequency there is a phase shift of $\phi = -\pi$ radians or -180 degrees. That means a change in sign. If the gain at this frequency is higher than one, the effect is similar to a positive feedback and the closed-loop system is unstable. If the gain is exactly one than a resonance occurs.

This observation can explain loop instabilities and forms the basis of the celebrated Bode and/or Nyquist stability criteria.

In the two integrator case, Fig. 2.13 without friction, we have exactly -180 degrees phase shift. Each integrator produces -90 degrees phase shift, so the total phase shift is $\phi = -180$. The gain in the loop is $1/\omega^2$, which is one for $\omega = 1$ (in this way we have identified the resonance pulsation).

For the three integrator problem, any constant feedback will result in an unstable system. To illustrate the effect of the feedback, we use a small trick. First we apply some small negative feedback loop around each of the integrators, so that we have a nice and stable behavior in the forward path. This is illustrated in Fig. 7.5. The phase shift for each of these systems varies from 0 to 90 degrees. Hence in the forward path the total phase shift varies from 0 for low frequency to 270 for high frequency inputs. (The phase shift of a linear systems consisting of two linear systems in series is indeed the sum of the phase shifts of the linear systems.)

Assume unitary feedback around the three-integrator system. The smaller the feedback gain we use internally in the components, the larger the gain in the forward path. This implies that at 180 degrees phase shift the loop gain can be larger than one, pro-

[6] A phase shift of $\pi/2$ or 90 degrees, a cosine wave input produces a sine wave output.

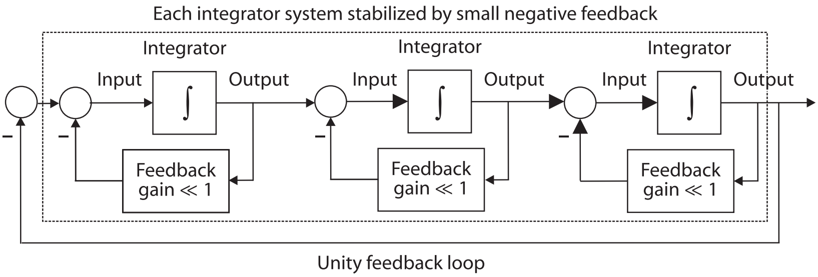

Fig. 7.5. Series connection of three feedback stabilized integrators, in a unity feedback loop; unstable for sufficiently small feedback gains

vided the introduced feedback gain is sufficiently small. Hence instability will result. The situation does not change when considering the limit, that is the case of zero feedback around the individual integrators, which is exactly the situation of three integrators in series in a negative feedback loop.

7.4.3 The Nyquist Stability Criterion

Proper design of feedback is a challenging problem in many real applications. One of the pioneers of feedback synthesis was Nyquist, who was faced with stabilizing telegraph telecommunication lines that required multiple signal amplifiers. He developed design rules, and the *Nyquist stability criterion* which settles the stability of single loop feedback systems for linear single input single output systems.

To capture the essence of this celebrated result, consider a unity negative feedback loop ($G_2 = 1$) with a plant in the forward path G_1 as in Fig. 5.8c, which is stable in itself. When is this loop stable? When is $(1 + G_1)^{-1}$ well-defined?

The problem can be appreciated as follows. Suppose that the input is a sinusoid. Because of linearity (and the fact that the plant is stable), the response is also sinusoidal with the same frequency, but with some phase shift and an amplitude change. In fact, all the signals in the loop will be sinusoidal in steady-state. If, for the applied frequency the open loop plant response is exactly 180 out of phase with the input with a gain of at least 1, then our negative feedback loop becomes a positive feedback loop. Because the gain is 1 or larger, chasing the signal around the loop indicates instability. If the gain was less than one, the loop results in some finite amplification, and all is well.

In general, the input of a system is not purely sinusoidal, but any signal can be decomposed in sinusoidal components (its power spectrum). Thus, if we analyze the frequency response of the plant (G_1), it will be easy to detect if for some frequency, the gain is larger than one and the phase shift being -180 degrees. If this is the case, the feedback system would be unstable. Graphically if we construct the (polar) plot of G_1 as a function of frequency, (an example in Fig. 7.6), in the complex plane the point $(-1, 0)$ must stay to the left, or remain outside the Nyquist contour[7].

[7] In the Nyquist criterion, a closed graph is plotted, the Nyquist contour, and the stability is determined by looking at the possible encircling of the (–1, 0) point.

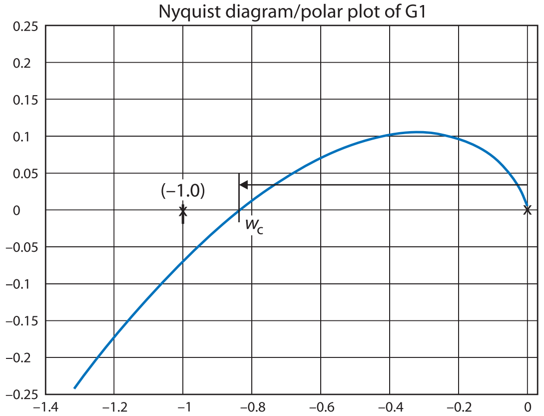

Fig. 7.6. Polar plot of a frequency response, from $\omega = 0$ (left) to $\omega \to \infty$ (origin). As the point $(-1, 0)$ is to the left of the curve, a unit feedback loop is well-defined

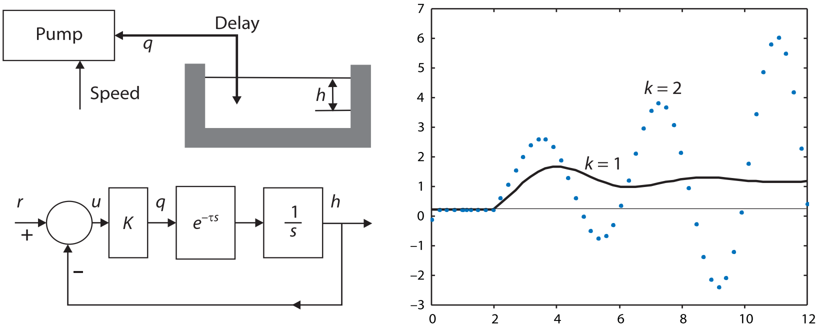

Fig. 7.7. Integrator with delay in a negative feedback loop

7.4.4 Integrator with Delay and Negative Feedback

Once the danger of positive feedback is appreciated, it is easy to see that systems with delay can cause havoc (like our shower!). A delay of τ seconds introduces a phase shift of value $\omega\tau$. Even a simple integrator with delay in a negative feedback loop can lead to instability when the gain is too large. This is illustrated in Fig. 7.7. Assume a simple tank, the net input flow q being controlled. The control signal is proportional, k, to the error in the tank level h with respect to a given reference r. It is not difficult to find when a gain is too large. The loop gain is k/ω. The total phase shift is $-(\omega + \pi/2)$. Therefore a phase shift of $-\pi$ occurs when $\omega = \pi/2$. The loop gain must be less than 1 at this frequency, so k is limited to be less than $\pi/2 \approx 1.5$. In the example in Fig. 7.7, a gain of 1 presents stable behavior, but a gain of 2 leads clearly to an unbounded response. In this case, the tank will either empty or overflow. Of course we will not see an unbounded signal, after all that would signify the end of the world as we know it.

7.5 Disturbance Rejection

Most systems are composed of many subsystems. This opens the possibility that in the connection between subsystems external, undesirable signals can enter. As the proper operation of the system depends on the information flow between systems, disturbances create problems.

Many different disturbances can be distinguished. In the case of signal transmissions they are in the main noise indicating some interference with or perhaps partial loss of signal. In other systems, as schematically represented in Fig. 6.11, the disturbances (signals d and n) may be more severe either corrupting a command signal or a measurement signal. Often it is either too difficult or too expensive to try to measure all possible disturbances. In this case, feedback can help to reduce the effect of these disturbances on the variables of interest. We say that feedback helps to make the system response *robust* with respect to the disturbances.

By way of example, consider an industrial boiler or steam generator (the central heating system in a building, or in a electrical power generating plant), as in Fig. 7.8, designed to provide steam (to heat, or to feed an engine).

In order to maximize the efficiency of the boiler it should work under design conditions of water level, temperature and pressure. If the boiler is going to run continuously with the same load, it may be enough to set some prescribed variables and let the system operate. Normally though the engine downstream or the building to be heated, does not need the same amount of steam all the time. The demand for steam is variable, and as a consequence the boiler should follow as to produce steam to meet demand. This has implications for feed water supply to the boiler and fuel supply to the burner. Steam demand is difficult to measure, and often somewhat unpredictable. Feedback comes to the rescue. Based on, for instance, the measurement of the boiler temperature, the fuel supply to the burner can be manipulated automatically. Feedback will react, if properly designed, in such a way that a decrement in the temperature will result in an increment of the fuel flow and vice versa. The influence of load variations will be greatly reduced and the temperature will remain close to the required value. Similarly the water level can be used to control the water supply, as shown in Fig. 7.8.

Again, by recalling the block diagram in Fig. 6.11, suppose that the disturbance d represents the steam demand, or load on the boiler. Without feedback, the variation in the

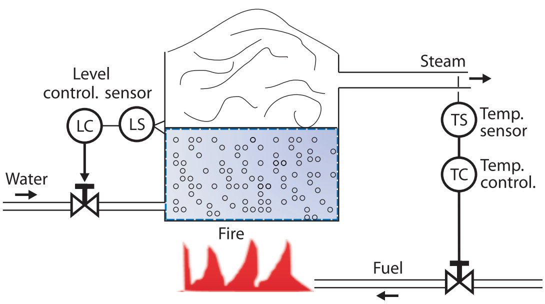

Fig. 7.8. Steam generator with some associated instrumentation

temperature (the output y) will be G times this disturbance. On the other hand, the relationship between the temperature and load variations when using the feedback will be

$$\frac{y}{d} = \frac{G}{1 + GKH} \tag{7.6}$$

and again, if the loop-gain of the composed operator GKH is much larger than one for the signals we are interested in this relation reduces to simply $1/KH$. The controller K can then be chosen to achieve a required disturbance rejection, i.e. K is large where d is large. For example if d were constant, an integrator in K will reject the load disturbance completely i.e. it is not necessary to know the amount of steam required. The boiler will deliver it!

7.5.1 Noise Feedback

One unavoidable aspect of feedback is that the system input will be influenced through the measurement on which the feedback acts. This implies that whenever this measurement has an error the feedback will produce erroneous activity. Any measurement implies some deviation between what had to be measured and what is measured, as no sensor is perfectly accurate. Say that the measured signal is the desired signal plus noise.

This noise will therefore excite the system unavoidably. This could lead to undesirable behavior. For example if measurement noise could excite the resonances in the antennae structure (see Sect. 3.4), the radio telescope would be a useless instrument. In Fig. 7.9, the speed of a motor is measured by a tachometer providing a noisy measurement.

Hence we need to measure as well as possible or, failing to do this, to filter the measurement as to suppress the noise and yet retain the necessary feedback information.

Filters, observers or estimators are different forms of elements in the feedback path that extract from the noisy measurement as much information about the system/signals as possible. Filters can exploit models of the system dynamics as an advantage, so as to be able to distinguish a signal that could come from the dynamics from another one which is external to the dynamics. Sometimes, the information that is extracted is related not only to the signal that is measured but also to other signals internal to the

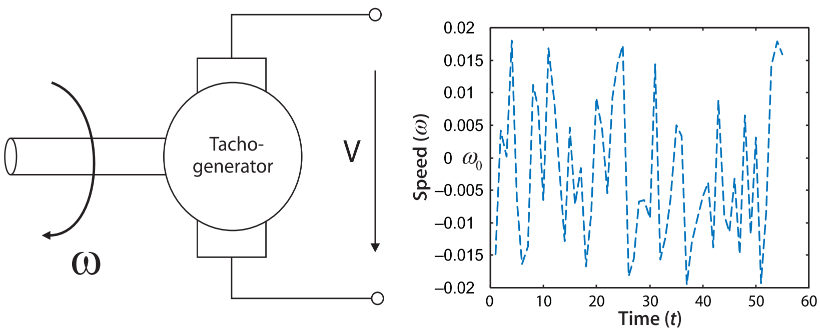

Fig. 7.9. A tachometer and the noisy signal it delivers

system, like the state components, as this may make it easier to compute what feedback is actually necessary. We will revisit this in the following chapters.

Looking at Fig. 6.11, the sensitivity functions previously defined (Eq. 6.17) are

$$T = S_{yr} = \frac{y}{r} = F\frac{GK}{1+HGK} \tag{7.7}$$

$$S = S_{yn} = \frac{y}{n} = \frac{1}{1+HGK}$$

Under the ideal conditions, ($HKG \gg 1$), these operators will be

$$T \approx \frac{F}{H} \quad ; \qquad S \simeq 0 \tag{7.8}$$

Thus, whenever the feedback filter H has a low gain at a given range of frequencies in order to reduce the effect of noise, the system output/reference will necessarily have high gain or amplify the input signals in this range of frequencies. This identifies another clear trade-off between requirements with respect to noise rejection and reference tracking.

Control Performance Trade-Off

There exists a fundamental trade-off in any feedback control system:

a high tracking accuracy inevitably reduces the capacity to reject (similar) disturbances.

Because typically the signals we want to track (steps, ramps, sinusoids) are quite different (at least from a frequency content point of view) from those appearing as output disturbances that must be rejected, this trade-off can be negotiated by "strategically" allocating tracking accuracy in the right frequency band.

Moreover, feedback control can be complemented with feed forward as in a two-degrees-of-freedom control strategy to improve overall performance (see also Sect. 8.6.1).

7.6 Two-Degrees-of-Freedom Control

If there are unwanted external signals, or disturbances acting on the plant, feedback can be designed to reduce their effect on the signals of interest. This happens without the need to measure the disturbance directly, and the feedback acts on the effect the disturbance has on the signal of interest (which is always measured, otherwise there would be no feedback). Well-tuned feedback reduces the sensitivity of the plant response to disturbances.

This property is very hard to achieve with feed-forward control, because to reduce the effect of a disturbance through the input to a cascade requires either precise measurement of the disturbance or precise (preview) knowledge. Also, in an open loop situation, if the systems in the cascade vary, this automatically affects the output. Using feedback it is possible to make some aspects of the response less dependent on these system variations, and thus maintain performance.

Consider again the typical control feedback system already discussed in the previous chapters (Fig. 6.11). The plant or system to be controlled is characterized by the operator G which is partially known, nonlinear, and perhaps time varying. The system is subject to an additive input disturbance d, which cannot be measured. For simplicity, let us assume $G_d = 1$. The plant output is affected by a (noise) disturbance n. The measurement device has a transfer operator H, which is used to suppress or filter out the effect of the disturbance n in the feedback loop. The feedback controller is K. A reference signal r is to be tracked by the output, i.e. we desire $y \approx r$.

From a design point of view, the questions revolve around the selection of K the feedback control, F the feed-forward control and to a lesser extent H, the measurement noise filter. The fact we have both K and F to work with, is captured as *two-degrees-of-freedom control*.

Feedback can be used to stabilize an unstable plant, attenuate disturbances, reject plant variations, and improve tracking performance. On the other hand, as already alluded to, too much feedback could destabilize the loop. Another unwanted effect is that feedback acts on what has been measured. So an incorrect measurement (as a consequence of the signal n in Fig. 6.11) will lead to unwanted control action.

Once feedback K (and H) has been selected mainly for responsiveness and disturbance rejection purposes, the tracking performance can be modified by means of F. As F is outside the loop, it does not influence the loop behavior too much, and does not affect stability at all. In this way two-degrees-of-freedom control can achieve both responsiveness and tracking performance. Because the feedback design makes the response less sensitive to plant variations, feedback is designed first, and then for the resulting compensated system the feedforward control is designed.

7.7 Feedback Design

Feedback may be used to get a controlled system that performs better than the open-loop system. Stability, tracking, disturbance rejection, are all aspects feedback may be used for. When one has the option of designing a new system that must have certain characteristics it pays to consider both the plant to be controlled and the feedback controller together to achieve the overall purpose. Of course, actually quite often, control arrives as an afterthought. The plant has been designed, and now some new requirements come forward, and feedback is added to an existing system. This has clear drawbacks.

Let us consider that you want to design a boat to sail in a race, fast and safe. First you could design a boat to be really safe, strong and able to master all winds and waves. This boat is presumably going to be rather heavy and hence sluggish. No amount of feedback used to optimally steer this boat, and trim the sails to the wind are going to make this boat behave like an America's Cup winner. More appropriately we would design a hull with a low keel, and little friction and large sails. The disadvantage is that wave action will be able to destabilize the boat, creating large oscillations, that would be virtually impossible to control by trimming the sails. Indeed this could not be guaranteed, as nobody can guarantee wind. Thus, creating a boat with an actuated keel, one that could be deployed and adjusted according to the wave action, and large sails that can be automatically trimmed would be the way to go. Not surprisingly the combined design of control system with plant to be controlled, leads to a superior solution.

Quite a similar situation may happen in designing the suspension of a car composed by dampers and springs. If they are passive elements, the stiffer they are the less displacement in the car body, but also the less comfort for the passenger. You may design them being very soft, but then large and sustained oscillations will result on a bumpy road. Feedback may provide the option to adjust the damping depending on the road conditions and the driving behavior: so-called active suspension.

In order to use feedback in systems' design as an advantage, the following conditions must be satisfied:

- well-defined goal(s) must be articulated;
- variables of interest must be measured, sensor subsystems must be provided;
- sensor information must be able to be related to goal attainment;
- the control algorithm must be able to decide how to take action, given information from the sensors and the goal(s), as well as the system model;
- feedback action can be applied as required to the system input, appropriately dimensioned actuators must be provided.

Open Loop Vs Closed Loop Control

Summarizing, open loop, as in a cascade or feed-forward control; closed-loop as in feedback control have the following characteristics:

- Open loop system controls cannot affect stability, those in closed loop can.
- Open loop systems are always well-posed, closed loop systems not necessarily.
- Open loop systems are not reactive, closed loop systems are.
- Open loop systems do not reject plant disturbances, closed loop systems can.
- Open loop systems are sensitive to plant variations, closed loop systems can suppress plant variations.
- Open loop control systems require accurate plant model and disturbance preview for performance, closed loop control systems do not rely on either an accurate plant model nor disturbance preview for performance.
- Open loop systems are not responsive to plant measurements, closed loop systems rely on accurate plant measurements.
- Combined feed-forward and feedback systems, two-degrees-of-freedom have all the advantages of either approach, and none of the disadvantages.

7.8 Discussion

Feedback is the most important feature in any controlled system.

Feedback is present in most natural as well as engineered systems. This follows because the basic building blocks in nature, as well as in the engineered world, are essentially very simple: accumulators (reservoirs, energy storage), gains, summers and delays. The behavior of these simple building blocks is nearly always trivial. The complex behavior so clearly observed in many systems, only follows through the interconnection of many simple subsystems using feedback loops, since without feedback, without creating loops, there is no complexity at all in the behavior.

Let us summarize the key ideas, and observations, together with some warnings:

- Feedback may be used to stabilize or destabilize a system.
- If the loop gain is less than unity, the closed loop is stable.
- Feedback may be used to improve robustness. To accomplish this high gain feedback is essential. If there is a high gain in the loop, the process model and or disturbances are less relevant and the global behavior mainly depends on the elements in the feedback path, which are specifically designed to achieve the control objective.
- High gain may result in components, like actuators, to saturate. High gain increases the risk for instability. There is a trade-off between robustness and stability.
- There is a trade-off between tracking and noise rejection.
- Two-degrees-of-freedom controllers help. The stabilization and disturbance rejection activities are tasks for the feedback loop, which is to be designed first. The tracking response is the task of the feed-forward controller, which can be designed next, based on the feedback stabilized loop.
- Feedback control reacts to errors detected in the system. Dealing with non-minimum phase systems (that is, systems showing an initial inverse response) or time delayed systems (that is, without any immediate response) require more sophisticated design. Simple high gain solutions will not work.

7.9 Comments and Further Reading

Feedback is a key concept in engineered systems and natural systems alike because of its ability to create interesting behavior from simple building blocks. As observed, feedback can significantly modify behavior. Feedback has a long history (Cruz and Kokotović 1972).

The importance of feedback in life as we know it, is well-established. See, for instance Hoagland and Dodson (1995), which gives a very gentle introduction. The importance of oscillations, and the role feedback plays to establish these in biology are well-documented in Goldbeter (1997).

A history of feedback control, with the emphasis on feedback starting with some interesting examples of clocks and float mechanisms is presented in Mayr (1970).

The input-output methodology is developed in Desoer and Vidyasagar (1975), and also in Mees (1991), which is heavily motivated by biology. Classical frequency domain ideas, building directly on the early work of Nyquist (1932) followed by Bode (1945) are developed in detail in Doyle et al. (1992) and a modern treatment using an algebraic and optimization-based approach is Green and Limebeer (1995). Despite the fact that a number of fundamental constraints have been established in feedback, we have only discussed superficially a few. No comprehensive or unifying treatment of feedback from this perspective exists in literature (even for linear systems).

A more mathematical treatment of feedback but still intended for a large audience, is Astrom and Murray (2008).

A treatise of feedback in a nonlinear systems and input-output setting is still under development. The role of input-to-state stability is developed in Sontag (1998) and Khalil (2002). Substantial advances have been made using ideas of passivity which we have not discussed. Passivity expands on the notion of system gain, and brings in the idea of phase, which plays such an important role in the linear system setting. A comprehensive treatment of this notion and how it applies to a very large class of nonlinear systems that can be described using the physical principles of electro-mechanical systems is Ortega et al. (1998). A more mathematical treatment is van der Schaft (2000).

The importance of feedback in electronic design is undisputed. A particular discussion of just this aspect of electronic design is Waldhauer (1982). The operational amplifier and its role in modern electronics is taught in all electrical engineering curricula.

The main control module used for feedback in the process industry is the so-called PID regulator, a good overview of the design questions surrounding PID controllers can be found in Astrom and Hagglund (2005).