A Short Introduction to Feedback Control

M.Z.

September 29, 2008

Stability is everything in the lab: Whether it is temperature, the current or frequency of a laser, the light-level coming through an optical fiber or the position of a mirror, we want to stabilize important parts of the experiment. This is achieved via feedback - you measure the quantity that you want to stabilize (say $x(t)$), compare it to a reference $r(t)$ (that you may choose to be time-dependent) and feed the difference signal back to your system, with the goal of making the quantity $x(t)$ converge as quickly as possible to the reference.

A mechanics analog Here comes a simple mechanics example by which I'll introduce the concepts of proportional, integral and differential feedback and how they work.

The model consists of an extended object of mass M, fully immersed in a stationary jar of honey and hanging from a spring with spring constant k and equilibrium length l (see Fig. 1). The spring is attached to a rod, that we hold in our hand to remotely act on the object. Gravity points down. We measure the position $x(t)$ of the object on a scale (on the left). The rod is at height $y(t)$. The viscous force due to the object's motion in the honey is $F_v = -\gamma\dot{x}$. Our goal is to force the object to follow as closely as possible and as quickly as possible a given reference height that we may choose to vary in time, $r(t)$. We want to achieve this by moving the rod in a certain fashion $y(t)$.

Let's first look at the total equation of motion of the object:

$$M\ddot{x} = -Mg - \gamma\dot{x} + k(y(t) - x(t) - l) \tag{1}$$

A transformation of the variable $y(t)$ into $y(t) + l + M/kg$ does away with

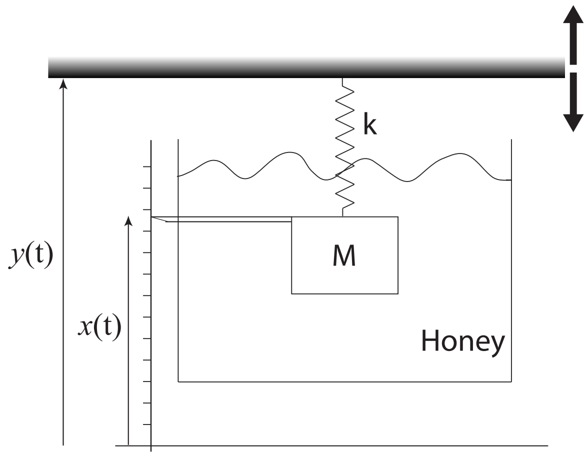

Figure 1: Feedback model 1: Mass on a spring in honey jar.

gravity and the equilibrium length, so we are left with:

$$\ddot{x} + \beta\dot{x} + \omega_0^2 x = \omega_0^2 y(t) \tag{2}$$

with $\omega_0 = \sqrt{k/M}$.

We can solve this equation for arbitrary $y(t)$. The easiest way is using the Fourier transform, $x(t) = \int \frac{\mathrm{d}\omega}{2\pi}\, e^{i\omega t} x(\omega)$. To demonstrate how to treat such problems in general, let's first solve for an "impulse", $y(t) = y_0\,\delta(t)$, whose Fourier transform is simply y_0.

$$\begin{aligned} -\omega^2 x(\omega) + i\,\beta\,\omega\,x(\omega) + \omega_0^2\,x(\omega) &= \omega_0^2 \cdot y_0 \\ x(\omega) &= y_0 \frac{\omega_0^2}{(\omega_0^2 - \omega^2) + i\beta\omega} \equiv y_0\,G(\omega) \end{aligned} \tag{3}$$

where we have defined the "impulse response function" $G(\omega)$:

$$G(\omega) = \frac{\omega_0^2}{(\omega_0^2 - \omega^2) + i\beta\omega} \tag{4}$$

The solution for an arbitrary drive $y(t)$ is then by superposition (since $y(t) = \int \mathrm{d}t'\delta(t-t')y(t')$) (we neglect transients)

$$x(\omega) = G(\omega)y(\omega) \tag{5}$$

Now we know how the object reacts to our drive. Our goal is for our object to track the reference height $r(t)$ faithfully. We could simply set $y(t) = r(t)$

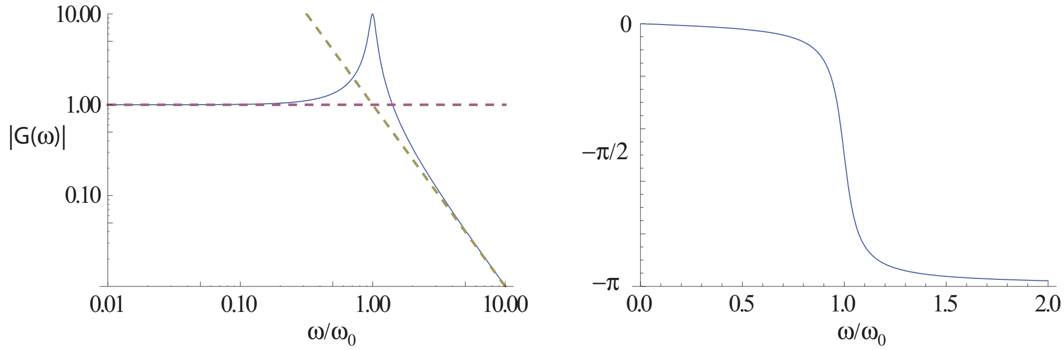

Figure 2: Amplitude response $|G(\omega)|$ and phase lag $\mathrm{Arg}(G(\omega))$ for the damped harmonic oscillator. These plots are called Bode magnitude and phase plots. The asymptotes 1 and ω_0^2/ω^2 cross at ω_0.

and indeed, for low frequencies $\omega \ll \omega_0$ we would have that $x(\omega) \approx r(\omega)$, the system would closely follow our drive. However, as we want our system to follow a drive at frequencies approaching ω_0, we realize that our object is too slow to react. Both amplitude and phase of the object's position are no longer those of the reference. On resonance, the amplitude becomes much larger than that of the reference, $|x(\omega_0)| = Q|r(\omega_0)|$, with $Q = \omega_0/\beta$ the quality factor of the oscillator. The phase lags behind by $\pi/2$. For higher frequencies, the amplitude decreases like $|x(\omega_0)| \to \frac{\omega_0^2}{\omega^2}|r(\omega)|$ and the phase lag approaches π (see Fig. 2).

For our system to track $r(t)$ more faithfully, we have to drive it with a more clever function $y(t)$, not $r(t)$ itself. Feedback will solve this problem for us - it "automatically" detects if the system lags behind or if its amplitude is too large, trying to make up for them. There is a second reason the simple drive by the reference itself is not enough. That reason is noise, which is always involved: For example our hand might not be very steady holding the rod, or there is wind randomly pulling and pushing at the rod or the object itself. All these effects add in as noise sources, so that in reality the drive is $y(t) = r(t) + \xi(t)$, with a random noise function $\xi(t)$. To make up for such random noise, we have to observe our system and try to cancel the noise by reacting to it, adapting our drive $y(t)$.

What we will have to do is to measure the current position of our object, $x(t)$, and compare it to the reference height $r(t)$ by defining an "error function"

$$\epsilon(t) = r(t) - x(t) \tag{6}$$

that measures the departure of the system from the desired path. The trick is now to chose our drive $y(t)$ as some linear functional K of $\epsilon(t)$. It is the exact form of K we actually want to optimize. So we try

$$y(t) = \int \mathrm{d}t' K(t-t')\epsilon(t') \tag{7}$$

A standard controller is the "PID"-control, which stands for proportional (P), integral (I) and derivative (D) gain:

$$y(t) = P\,\epsilon(t) + I\int_0^t \mathrm{d}t'\epsilon(t') + D\,\dot{\epsilon}(t) \tag{8}$$

In Fourier space:

$$y(\omega) = K(\omega)\,\epsilon(\omega) \tag{9}$$

$$y(\omega) = \left(P - i\frac{I}{\omega} + iD\,\omega\right)\epsilon(\omega) \tag{10}$$

Now we can "close the loop", i.e. plug this drive y, that is now dependent on the deviation $\epsilon = r - x$, back into the dynamical equation for x:

$$x(\omega) = G(\omega)y(\omega) = G(\omega)K(\omega)\epsilon(\omega) = G(\omega)K(\omega)(r(\omega) - x(\omega)) \tag{11}$$

We see how we now have $x(\omega)$ on both sides of the equation. The solution is

$$x(\omega) = \frac{G(\omega)K(\omega)}{1 + G(\omega)K(\omega)}\,r(\omega) \tag{12}$$

Plugging in the known impulse response function for our system, we get

$$x(\omega) = \frac{K(\omega)}{(1 - \frac{\omega^2}{\omega_0^2}) + i\beta\frac{\omega}{\omega_0} + K(\omega)}\,r(\omega) \tag{13}$$

Via feedback, we have modified our original relation Eq. 5 for the "bare" system into a different relation that we can tweak to make our system follow the given path $r(t)$ much more faithfully. Note that we have chosen $\epsilon = r - x$

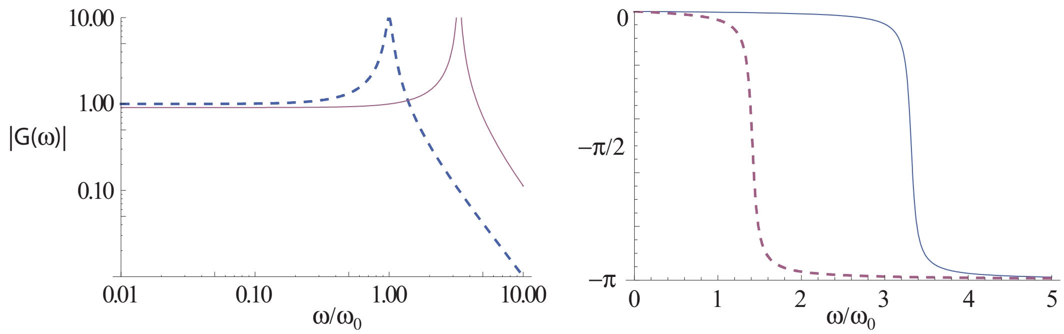

Figure 3: Amplitude response $|F(\omega)|$ and phase lag $\mathrm{Arg}(F(\omega))$ for the damped harmonic oscillator with proportional feedback. The "P"-gain increased the resonance frequency of the "closed loop"-system by $\sqrt{1+P}$.

with a minus sign in front of the measurement x. If x is too low, this gives a positive error, if it is too high, ϵ is negative. Thus, the rror always "looks in the direction" in which we can decrease the error. A simple proportional feedback $y = P(r - x)$ with $P > 0$ will generally tend to drive the system back towards the reference. Let's see how this works in the time domain:

$$\ddot{x} + \beta\dot{x} + \omega_0^2 x = \omega_0^2 y(t) = \omega_0^2 P(r(t) - x(t)) \tag{14}$$

Rearranging gives:

$$\ddot{x} + \beta\dot{x} + \omega_0^2(1 + P)x = \omega_0^2 P r(t) \tag{15}$$

This describes a driven harmonic oscillator with a much higher resonance frequency (if $P \gg 1$) $\omega_P = \omega_0\sqrt{1+P}$. Our oscillator can thus follow the reference without appreciable phase lag for much faster variations than the "bare" system. The amplitude relation is

$$x(\omega \ll \omega_P) \approx \frac{P}{1+P} r(\omega) \approx r(\omega) \qquad \text{if } P \gg 1 \tag{16}$$

so we are quite successful in imposing our reference r onto the system (see Fig. 3).

One price to pay is that the resonance at ω_p has now a higher quality, $Q_p = Q\frac{P}{\sqrt{1+P}}$. Thus, if there is noise around that frequency, it will be resonantly enhanced. In fact let us distinguish two types of noise: Disturbances of the object or the rod itself, $d(t)$ and an error in our measurement of the object's position, $\xi(t)$. The total dynamics of our object will then be described by

$$x = KG\epsilon + d \tag{17}$$

with

$$\epsilon = r - x - \xi \tag{18}$$

This implies that

$$x = \frac{KG}{1+KG}(r-\xi) + \frac{1}{1+KG}d \tag{19}$$

(You can read the equation either as an operator equation or as it's realization by direct multiplication in Fourier space). As we see, the disturbances d are suppressed up to a frequency ω_p, which is what we want. However, the system has no way to distinguish the true, wanted reference r from the measurement noise ξ. Thus, the higher the gain, the more measurement noise we couple in. This can be expressed in the "tracking error" ϵ_0, the difference between the desired (error free) reference r and the actual position of the object x:

$$\epsilon_0 \equiv r - x = S(r-d) + T\xi \tag{20}$$

with the "sensitivity function" $S = \frac{KG}{1+KG}$ and the "complementary sensitivity function" $T = 1 - S = \frac{1}{1+KG}$. The overall goal is to have ϵ_0 be as small as possible. A fundamental obstacle is that $T + S = 1$ at all frequencies. Thus, for small S disturbances are rejected but then T is large and measurement noise is coupled in, and vice versa.

Integral gain The proportional gain, although successful in suppressing noise even at frequencies around ω_0, suffers from "proportional droop": The dc-value $r(\omega = 0)$ is not faithfully reproduced, as $x(\omega = 0) = \frac{P}{1+P}r(\omega = 0)$. This can be taken care of by integral gain. The error signal is integrated, $I\int_0^t(r_0 - x(t))$, so that the longer the system is not at the set height r_0, the more the object is pushed towards r_0. The equation of motion reads:

$$\ddot{x} + \beta\dot{x} + \omega_0^2 x = \omega_0^2 I\int_0^t (r_0 - x(t)) \tag{21}$$

Differentiating, we get

$$\dddot{x} + \beta\ddot{x} + \omega_0^2\dot{x} = \omega_0^2 I(r_0 - x(t)) \tag{22}$$

and we see that in steady state we will have $x(t) = r_0$. We can also see this in the frequency domain, where $K(\omega) = -i\frac{I}{\omega}$. This can be thought of as frequency dependent gain, and there is infinite gain at dc ($\omega \to 0$).

Instability Looking again at the solution for the system with feedback:

$$x(\omega) = \frac{G(\omega)K(\omega)}{1 + G(\omega)K(\omega)}\, r(\omega) \tag{23}$$

we realize that there is an instability if $G(\omega)K(\omega) = -1$. $x(\omega)$ would grow to infinity, signalling an exponentially growing mode at that frequency. Since G and K are complex, we have two conditions for instability, $|G||K| = 1$ and the total "open loop" gain GK should have a phase lag of 180 degrees. For our bare system G, this does not seem to be an issue: The second derivative $\ddot{x}$ implies a term $(i)^2\omega^2 = -\omega^2$ in the denominator of $G(\omega)$, signalling that the 180 phase shift is only reached at high frequencies $\omega \gg \omega_0$. Proportional gain $K = P\epsilon$ does not change this fact, so we do not seem to run into instabilities with purely a P-gain (this is not true in an actual system, where there is always a time lag between the measurement and the drive, see below). With integral gain, the situation gets worse, because $K(\omega) = -iI/\omega$ adds another 90 degrees phase lag to GK. With purely integral gain, right at the bare resonance $\omega = \omega_0$ we would have an overall phase lag of 180 degrees instead of 90. At the same time, we must have $|G||K| < 1$ or else we will have an unstable mode. This would limit our integral gain to $I < \omega_0/Q$ (see Fig. 4).

If we switch on proportional gain in addition to integral gain, the situation is much better (see Fig. 5). The 180 degrees phase lag is avoided while still having infinite dc gain. However, in an actual system, there is always a finite delay between the measurement and the drive. A constant delay for example would yield a drive $y(t) = \epsilon(t - \tau)$, which in Fourier space reads $y(\omega) = e^{-i\omega\tau}\epsilon(\omega)$. This adds a phase lag that is proportional to frequency! Thus, eventually any system will cross the 180 degrees phase lag. It is at this point we need to make sure $|K||G| < 1$ to prevent instabilities.

Derivative gain It seems obvious it will help to include a derivative term in K, setting $K(\omega) = iD\omega$. This advances the phase at high frequencies and

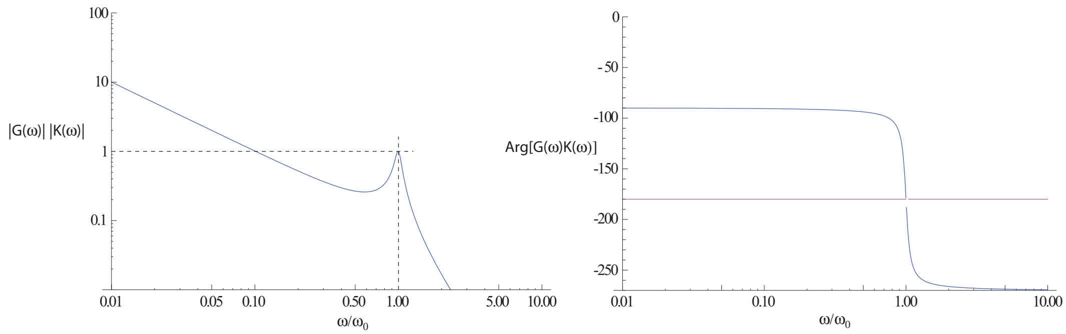

Figure 4: Pure integral (I) gain gives infinite feedback at dc, but suffers from a phase lag of -90 degrees. For our harmonic oscillator system with a resonance at ω_0, we need to limit the gain on resonance, where the phase lag is 180 degrees and we would have positive instead of negative feedback.

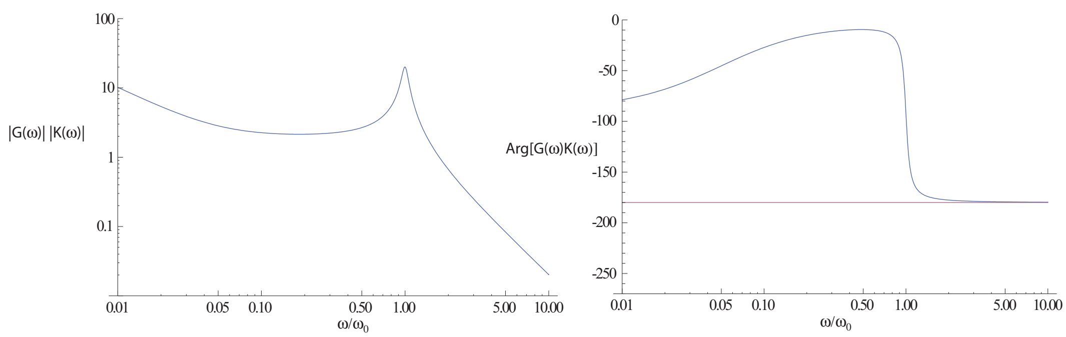

Figure 5: Proportional and integral gain gives still infinite feedback at dc and it we do not reach a phase lag of 180 degrees anymore.

helps to prevent the 180 degrees phase lag. This means that if we observe the object $x(t)$ to fall down at rapid speed, this must be due to some strong perturbation that we want to counteract immediately, not waiting for proportional or integral gain until $x(t)$ has dropped below the set point. One obvious drawback of derivative gain is that it will increase the importance of measurement noise at high frequencies, which will now be coupled in effectively. One simple solution to this is to "roll off" the derivative feedback above some cut-off frequency ω_1, thus setting $K(\omega) = iD\omega/(1 + i\omega/\omega_1)$.

PID-Controller The total PID-gain is $K(\omega) = P - i\frac{I}{\omega} + iD\omega$. As it turns out, by tweaking the parameters right, we can exactly cancel the effect of the resonance in $G(\omega) = \omega_0^2/(\omega_0^2 - \omega^2 + i\beta\omega)$ (see Fig. 6).

$$\begin{aligned}
K(\omega)G(\omega) &= \omega_0^2 \frac{P - i\frac{I}{\omega} + iD\omega}{\omega_0^2 - \omega^2 + i\beta\omega} \\
&= \frac{-iD\omega_0^2}{\omega} \frac{\frac{I}{D} - \omega^2 + i\frac{P}{D}\omega}{\omega_0^2 - \omega^2 + i\beta\omega} \\
&= \frac{-iI}{\omega} \\
\text{with } \frac{I}{D} &= \omega_0^2, \qquad \frac{P}{D} = \beta \qquad (24)
\end{aligned}$$

Instead of a second order system (one with $(i\omega)^2$ in the denominator), the open-loop gain is now first order (a simple integrator), and the closed loop gain reads

$$T(\omega) = \frac{KG}{1 + KG} = \frac{1}{1 + i\omega/I} \qquad (25)$$

which is a simple low-pass filter with time constant $\tau = 1/I$.

The experimental "tuning" of a PID-controller is shown in Fig. 7. Using a step function as the reference one first starts with the I-gain until the system runs into oscillations. Backing up a bit, one increases P (which in this idealized system cannot lead to ever growing oscillations), then D, which usually makes the feedback faster and reduces oscillations, as the phase lag is reduced from 180 degrees at high frequencies. For the proper choice of the P,I, D-values, one should observe that the system behaves as a low-pass, without oscillations (critical damping of oscillations).

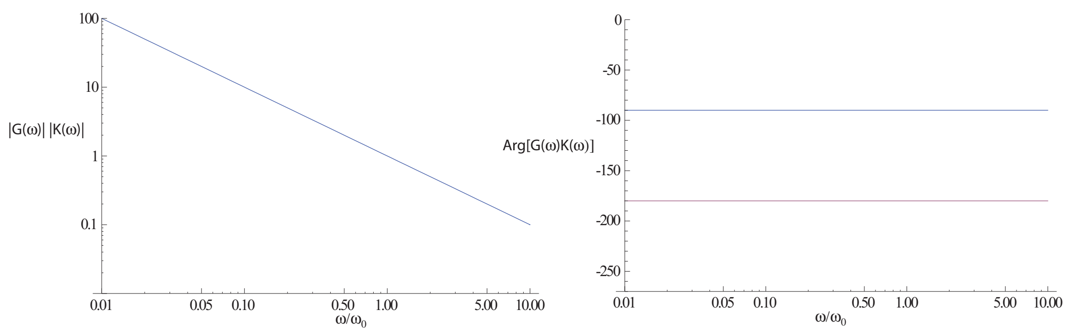

Figure 6: PID-gain. The resonance of the original system has been "tweaked away" by proper choice of the PID-parameters.

Other examples The techniques described above work for all kinds of linear systems (where the equation of motion is linear in the variables). To give another mechanics example, we could, instead of pulling the rod, move the jar up and down to control the relative velocity of honey and object. Thus, we control the system via the viscous force, not via the spring force (see Fig. 8).

The equation of motion reads:

$$\ddot{x} + \beta(\dot{x} - \dot{y}) + \omega_0^2 x = 0 \tag{26}$$

The bare impulse response function is the same as before, but the feedback is now via $\dot{y}$.

We often have to stabilize the temperature of a laser diode using a thermoelectric cooler (TEC) as heating/cooling element and a thermistor as temperature sensor. A first order approximation to this system is to say that the current / power delivered to the TEC will modify the temperature of the system according to

$$\frac{dT}{dt} = \frac{1}{C}\frac{dQ}{dt} = -\frac{1}{C}W_{\text{out}} + \frac{1}{C}W_{\text{in}} \tag{27}$$

Usually, $W_{\text{out}} = \alpha A \Delta T$ with α a material-dependent coefficient, A the surface area of the block that we want to stabilize, and ΔT the difference of the block's temperature and the ambient temperature. We thus get an equation of motion

$$\dot{\Delta T} + \frac{\alpha A}{C}\Delta T = \frac{1}{C}W_{\text{in}} \tag{28}$$

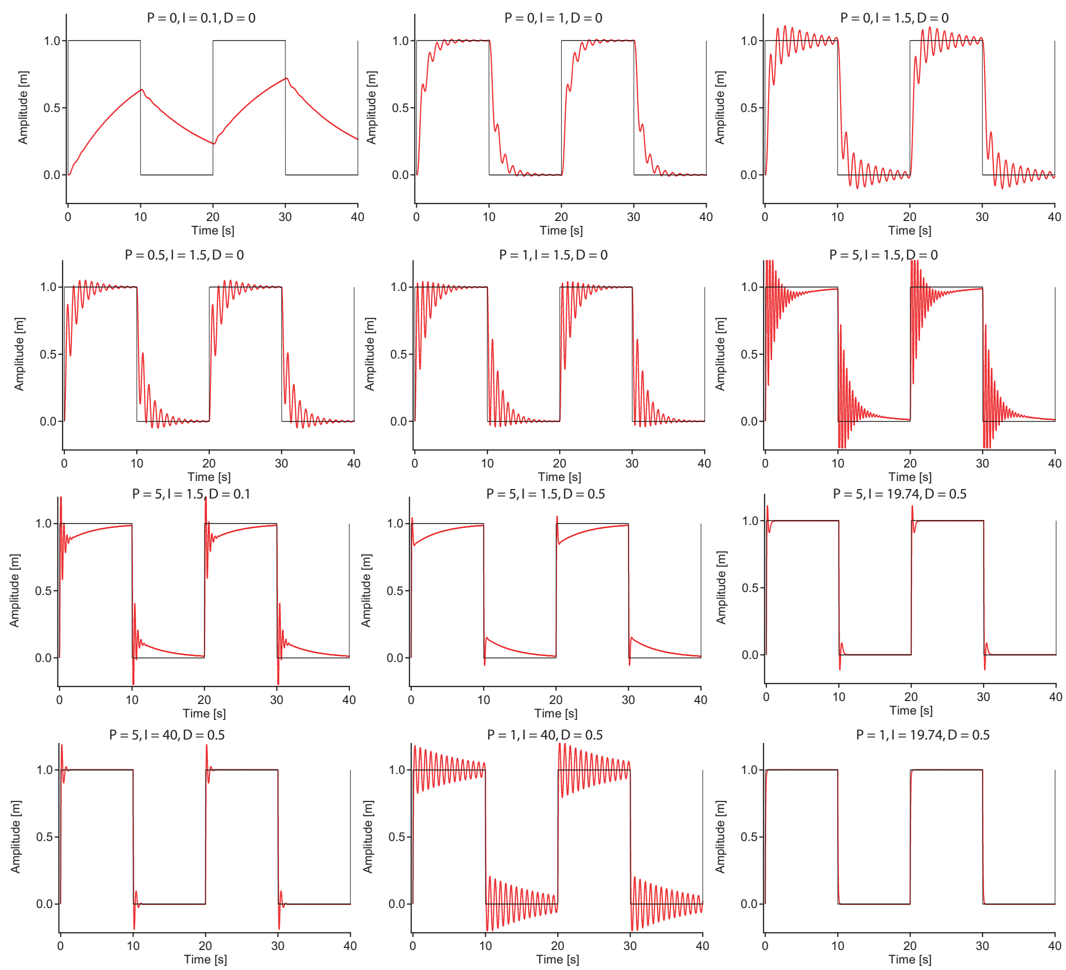

Figure 7: PID-control: The goal is to have the object follow the reference path of $r(t)$ (black), which here is a step function. The bare oscillator has $\omega_0 = 2\pi \cdot 1\,\mathrm{Hz}$, and $\beta = 2\frac{1}{\mathrm{s}}$. In the first row integral gain is increased until oscillations (of period $2\pi/\omega_0$!) are enhanced. In the next row proportional gain P is added. Note how the visible oscillations increase in frequency, as the system starts to act like an oscillator with $\omega_p = \sqrt{1+P}\,\omega_0$. Increasing P does not lead to an instability in this idealized case without time delays. In the third row derivative (D) gain is switched on, which helps to reduce the strong oscillations due to the smaller phase lag. The I-gain is now increased first to $\approx$ 201/s, then to 401/s (fourth row). A reduction of the P-gain brings back some oscillations, which are reduced to a minimum by setting $I = 19.741/s = D\omega_0^2$. (last graph). Here, we see how the system behaves as a low-pass filter, no oscillations are present. The optimum P,I,D values have been found.

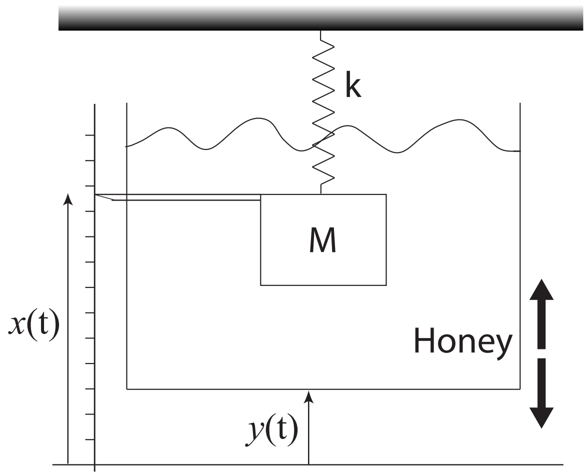

Figure 8: Feedback model 2: Mass on a spring in honey jar. This time, the velocity of the jar itself acts as the controller.

The bare system is simply a low-pass filter, with time constant $\tau = C/\alpha A$, which one can measure by simply heating the (bare) system by about a degree and observing the decay back to room temperature. In Fourier space, we have

$$\Delta T = \frac{1}{1 + i\omega\tau}\frac{W\tau}{C} \equiv G\frac{W\tau}{C} \tag{29}$$

We now switch on feedback by setting $W\frac{\tau}{C} = K\epsilon$ with the error function $\epsilon = r - \Delta T$ where r is the reference temperature we want to reach (possibly time-dependent, but typically not). Let's switch on a proportional gain, $K = P$. This gives

$$\Delta T = \frac{KG}{1 + KG}r = \frac{P}{1 + i\omega\tau + P} = \frac{P}{1 + P}\frac{1}{1 + i\omega\frac{\tau}{1+P}}r \tag{30}$$

This is again a low-pass filter, but with modified gain $\frac{P}{1+P}$ and faster response (for $P > 1$), namely time constant $\tau/(1 + P)$. This means we would like infinite P-gain. However, this assumes that we truly read the temperature of the block right at the laser diode we want to stabilize. In reality, the thermistor is a certain distance away from the diode, and it will intrinsically not respond immediately to the surrounding temperature. Also, the TEC itself is located not at the laser diode but beneath the block on which it sits, so again, there will be a time lag between the moment we apply the power and the temperature change of the diode. We have seen before how such time lags lead to phase lags that grow linearly in frequency. The propagation time of heat transfer through the block depends on the thermal conductivity λ, the density ρ and the heat capacity C of the block. For a distance L between TEC cooler and laser diode, the time lag is L^2/D with $D = \lambda/\rho C$ the diffusion constant of the block material. Thus, at a characteristic frequency D/L^2, the power at the TEC element and the temperature at the diode will be 180 degrees out of phase. At this frequency, we will need to have the gain $|K||G| < 1$.

Note that including a derivative (D)-gain, we could cancel the low-pass behavior altogether by setting $D/P = \tau$.

With only P and D-gain, we again have the problem of "proportional droop", the temperature settles to $\frac{P}{1+P}$ times the reference. Adding integral gain solves this problem by providing infinite gain at dc. P- and I-gain gives

$$\begin{aligned} KG &= \frac{P - \frac{i}{\omega}I}{1+i\omega\tau} = -\frac{iI}{\omega}\frac{1+i\omega\frac{P}{I}}{1+i\omega\tau} \\ &= -\frac{iI}{\omega} \qquad \text{with } \frac{P}{I} = \tau \end{aligned} \tag{31}$$

The choice $P/I = \tau$ results in an open-loop gain of I/ω and a phase lag of 90 degrees. The closed-loop gain is $T = \frac{KG}{1+KG} = \frac{1}{1+i\omega/I}$, a simple low-pass with band-width I. In truth, we need to make sure that $I < D/L^2$, to avoid oscillations due to the finite heat diffusion time.

A simple way to measure the phase lags of the bare system is to drive it with a sinusoidal power, using a frequency generator as the current controller, and observing the thermistor-readout together with the drive on a scope. This directly allows to read off phase lags.

References

1. Bechhoefer, J., "Feedback for Physicists: A tutorial essay on control", Rev. Mod. Phys. **77** (2005).

2. Forgan, E.M., "On the use of temperature controllers in cryogenics", Cryogenics **14**, 207 (1974).