Balancing a stick

The parameters of the system are:

Parameter	Definition
m	mass of the stick, unknown
g	acceleration of gravity, (9.81 m/s^2)
L	height of the center of mass of the stick from the pivot (0.63 m)
L_0	height of the fixation point measured from the pivot (1.0 m)
J	moment of inertia of the stick with respect to the pivot, unknown

The variables of interest are:

Variable	Definition
θ	angle of the stick measured with respect to vertical
x	horizontal position of the pivot (hand)
z	horizontal position of the fixation point
w	sensor noise (in the eyes)
y	noisy measurement of horizontal position of fixation point
u	input (acceleration of the hand/pivot)

Taking a torque balance about the pivot and using a small-angle approximation, we obtain the equation of motion:

$$J\ddot{\theta} + mL\ddot{x} = mgL\theta$$

We also assume the input is the hand acceleration, so

$$u = \ddot{x}$$

The eye observes the lateral position at a height L_0 above the pivot, which is given by (again using small-angle approximation):

$$z = x + L_0\theta$$

However, our eyes aren't perfect, and there will be measurement noise. So the true signal used by our controller is

$$y = z + w$$

Transfer function

Although we don't know J, it must be some multiply of mL^2. For example, if the rod were a point mass at L, we would have $J = mL^2$. If we had a uniform rod with center of mass at L (total length $2L$), then we would have $J = \frac{1}{3}m(2L)^2 = \frac{4}{3}mL^2$. Let's define $J := \eta m L^2$ from now on.

$$\begin{aligned}(\eta L s^2 - g)\Theta + s^2 X &= 0 \\ s^2 X &= U \\ Z &= X + L_0 \Theta \\ Y &= Z + W\end{aligned}$$

Eliminating X, Z, and Θ, we obtain the equation:

$$Y = \left(\frac{(\eta L - L_0)s^2 - g}{s^2(\eta L s^2 - g)} \right) U + W$$

Divide numerator and denominator by ηL:

$$Y = \left(\frac{(1 - \frac{L_0}{\eta L})s^2 - \frac{g}{\eta L}}{s^2(s^2 - \frac{g}{\eta L})} \right) U + W$$

Define $\omega_n^2 = \frac{g}{\eta L}$ and we can write:

$$Y = \left(\frac{(1 - \frac{L_0}{\eta L})s^2 - \omega_n^2}{s^2(s^2 - \omega_n^2)} \right) U + W$$

Measuring parameters

We do not need to find J or m since our equation depends only on (ω_n, L, L_0, η).

Note that the original unforced pendulum has equation $\ddot{\theta} - \omega_n^2 \theta = 0$. If the pendulum were not inverted, then its equation would be $\ddot{\theta} + \omega_n^2 \theta = 0$. So ω_n is indeed the natural frequency. We can measure ω_n directly, by simply suspending the rod from a string and timing the period, and we find $T = 1.8$ seconds. The frequency is:

$$\omega_n = \frac{2\pi}{T} \approx 3.49 \text{ rad/sec}$$

Recalling the definitinon $\omega_n^2 = \frac{g}{\eta L}$, we know that $g = 9.81 \text{ m/s}^2$ and we can measure $L = 0.63 \text{ m}$ by balancing the rod to find the center of mass. Solving for η, we find:

$$\eta = \frac{g}{\omega_n^2 L} = \frac{9.81 \text{ m/s}^2}{(3.49 \text{ rad/s})^2 (0.63 \text{ m})} \approx 1.28$$

So we now have a complete description of the equations of motion.

Analysis of poles and zeros

The transfer function from U to Y has four poles: $\{0, 0, \pm\omega_n\}$, where $\omega_n \approx \pm 3.49$ rad/sec. We can identify three cases for the zeros:

$$\begin{cases} s = \pm j \frac{\omega_n}{\sqrt{\frac{L_0}{\eta L} - 1}} & \text{if } L_0 > \eta L \approx 0.8 \\ \text{no zeros} & \text{if } L_0 = \eta L \approx 0.8 \\ s = \pm \frac{\omega_n}{\sqrt{1 - \frac{L_0}{\eta L}}} & \text{if } L_0 < \eta L \approx 0.8 \end{cases}$$

Drawing the root loci for these three cases, we can see that having imaginary zeros is favorable. The case with a RHP zero seems particularly difficult the smaller L_0 gets, because this means the RHP zero is very close to the unstable pole, which will make it quite difficult to attract the root locus to the left-half plane.

Complementary sensitivity

We are interested in how the noise will be amplified. Let's call G the transfer function from U to Y, so $Y = GU + W$. If the controller is K, then we have $U = K(R - Y)$, where R is the reference. In our case, $R = 0$, so we have $U = -KY$. Substituting into above and solving for Y, we find:

$$Y = \frac{1}{1 + GK} W$$

The signal $R - Y$ is what the controller sees, but it's not the true error, since it includes sensor noise. The true error is $E = R - Z$. We find this to be:

$$E = R - Z = R - Y + W = 0 - \frac{1}{1 + GK} W + W = \underbrace{\left(\frac{GK}{1 + GK} \right)}_{T} W$$

The transfer function T is our familiar closed-loop map. It's also the transfer function from sensor noise to true error. It also has the name *complementary sensitivity function*.

Fundamental limits

We don't know what K is; it's the human body and it's complicated. However, there are fundamental limits to controller performance, and these limits cannot be overcome. These are essentially laws of nature, like conservation of energy, or the laws of thermodynamics. We will learn about some of these limits now.

Ideally, we would want $|T(j\omega)|$ to be small for all ω, since this would mean that every frequency of noise is weakly amplified and does not appear prominently in the error. But is it possible to make it arbitrarily small? Let's call M_T the maximum magnitude of T. So we're defining:

$$M_T := \max_{\omega} |T(j\omega)|$$

We will make use of a result called the *maximum modulus theorem*. It says that if $H(s)$ is any transfer function that is bounded in the entire right-half plane (including the imaginary axis), then its maximum magnitude must occur somewhere *on* the imaginary axis. Here, "bounded" means that we can't make the magnitude arbitrarily large. In particular, it means that H must be stable and can't have any poles on the imaginary axis either. If this were not the case, then picking s close to the pole in the right-half plane would make the magnitude blow up. This also means H should be proper, so that as $\mathrm{Re}(s) \to \infty$ we should have $H(s) \to 0$.

Since K should stabilize our system, T must be stable. So we can apply the maximum modulus theorem to T. Note also that G has an unstable pole at $s = \omega_n$. We're assuming that K cannot contain a RHP zero at ω_n (remember: it's dangerous to try and cancel unstable poles using RHP zeros!). Therefore, $T(j\omega_n) = 1$. Applying the maximum modulus theorem, we have:

$$M_T = \max_{\omega} |T(j\omega)| = \max_{s \in \mathrm{RHP}} |T(s)| \geq |T(\omega_n)| = 1$$

Here, "RHP" denotes the entire right-half plane including the boundary: $\{s \in \mathbb{C} \mid \mathrm{Re}(s) \geq 0\}$. So no matter what we do, the maximum noise amplification will be at least 1 (this is guaranteed to occur at $s = \omega_n$). *It is impossible to attain attenuation at all frequencies, no matter what controller is used*. This isn't the end of the story though...

Time delay

The human body has a natural time delay in vision + processing + motor control. This loop has a delay of about 300 milliseconds (0.3 seconds) in most humans. This has consequences for noise amplification, as we will now see.

Assuming our controller also carries a delay of $\tau = 0.3$ seconds, that means we can multiply K by $e^{s\tau}$ and it will remain bounded in the RHP. An ordinary (non-delayed) transfer function could not tolerate this, as it would blow up as $|s| \to \infty$ because exponentials grow faster than any polynomial. Also, exponentials always have magnitude 1. So using the maximum modulus theorem, we have:

$$M_T = \max_{\omega} |T(j\omega)| = \max_{\omega} |e^{j\omega\tau} T(j\omega)| = \max_{s \in \mathrm{RHP}} |e^{s\tau} T(s)| \geq |e^{\omega_n \tau} T(\omega_n)| = e^{\omega_n \tau}$$

For our system, $\omega_n = 3.49$ and $\tau = 0.3$. Therefore, we have:

$$M_T \geq e^{\omega_n \tau} \approx 2.85 \approx 9 \text{ dB}$$

So in the presence of delay, the noise amplification is at least 9 dB, and this is unavoidable!

Right-half plane (RHP) zeros

We saw above that if $L_0 < \eta L \approx 0.8$, we will have a RHP zero. This also has consequences! If there is a RHP (real) zero at $s = z$, then G has a factor $(s - z)$. Consider the unstable filter with transfer function

$$F(s) = \frac{s+z}{s-z}$$

First, observe that $|F(j\omega)| = 1$ for all ω. This follows because:

$$|F(j\omega)| = \left|\frac{j\omega + z}{j\omega - z}\right| = \frac{\sqrt{\omega^2 + z^2}}{\sqrt{\omega^2 + z^2}} = 1$$

Next, notice that even though F is unstable, FG is stable, because z is a zero of G, so it gets canceled by multiplication with F. Therefore, we can apply maximum modulus again and write:

$$\begin{aligned} M_T &= \max_{\omega} |T(j\omega)| = \max_{\omega} |F(j\omega)T(j\omega)| \\ &= \max_{s \in \text{RHP}} |F(s)T(s)| \geq |F(\omega_n)T(\omega_n)| = \left|\frac{\omega_n + z}{\omega_n - z}\right| \end{aligned}$$

The RHP zeros is located at $z = \frac{\omega_n}{\sqrt{1 - \frac{L_0}{\eta L}}}$. Substituting, we obtain:

$$M_T \geq \left|\frac{1 + \sqrt{1 - \frac{L_0}{\eta L}}}{1 - \sqrt{1 - \frac{L_0}{\eta L}}}\right|$$

Recall $\eta L \approx 0.8$. Let's assume we fix our gaze on the midpoint of the stick. Then $L_0 = 0.5$, and

$$M_T \geq \left|\frac{1 + \sqrt{1 - \frac{5}{8}}}{1 - \sqrt{1 - \frac{5}{8}}}\right| \approx 4.16 \approx 12.4 \text{ dB}$$

In fact, we still have the delay, so the total amplification is $2.85 \times 4.16 = 11.86$, or $9 + 12.4 = 21.5$ dB.

Shortening the stick

If we shorten the stick, say to half the length, then let's assume η doesn't change, and L and L_0 each get cut in half. Since $\omega_n \propto \frac{1}{\sqrt{L}}$, the new natural frequency will be $\omega_n' = 3.49 \times \sqrt{2} \approx 4.94$. This will cause the noise amplification due to delay to become:

$$M_T \geq e^{\omega_n' \tau} \approx 4.4 \approx 12.8 \text{ dB}$$

So shortening the stick incurs an extra 4 dB of amplification, but lowering our gaze to the half-point incurs an extra 12.4 dB, which is much worse!

Bandwidth considerations

Not only is it impossible to avoid noise amplification at certain frequencies, it is also impossible to avoid a high closed-loop bandwidth. To reason about this, let's suppose we would like to achieve a closed-loop transfer function that is at least as good as a first-order low-pass filter. In other words,

$$|T_{\text{goal}}(j\omega)| = \left|\frac{M_T}{\frac{j\omega}{\omega_c}+1}\right| \geq |T(j\omega)|. \qquad \text{Therefore,} \quad 1 \geq \left|\frac{T(j\omega)}{T_{\text{goal}}(j\omega)}\right|.$$

Here, ω_c is the corner frequency for our desired closed-loop map and M_T is the maximum amplification. Let's apply maximum modulus to the function T/T_{goal} and substitute the unstable pole as we did before:

$$1 \geq \max_{\omega}\left|\frac{T(j\omega)}{T_{\text{goal}}(j\omega)}\right| = \max_{s\in\text{RHP}}\left|\frac{T(s)\cdot(s/\omega_c+1)}{M_T}\right| \geq \frac{\omega_n/\omega_c+1}{M_T}$$

Rearranging, we obtain:

$$(M_T-1)\omega_c \geq \omega_n$$

- We want ω_c to be small; this means the closed-loop bandwidth is small, which means we reject a lot of the noise, particularly at high frequencies.
- We also want M_T to be small (recall it has to be at least 1), because this means that the noise frequencies that are most amplified are not amplified that much.

The inequality above tells us that ω_c and M_T can't both be small. Moreover, the smaller we make L, the larger ω_n will be, and the harder it becomes to satisfy our design requirement. This is a fundamental *trade-off* in controller design. It's also sometimes called the *waterbed effect* (you can push some frequencies down, but others will pop up).

We can also combine all aforementioned effects together. If $M_T \geq M_0$ due to a RHP zero and/or a time delay, then applying a similar argument to above, we obtain

$$(M_T-M_0)\omega_c \geq M_0\omega_n$$

Misc.

- Control gets virtually impossible if one eye is closed, because we lose depth perception and we can no longer estimate the distance of the rod in the fore-aft direction.
- Adding more weight to the tip makes control easier, because this moves the center of mass higher (increases L). Note that the total mass makes no difference, only the mass distribution is important.
- Adding sensor noise by dimming the lights in the room or standing on one leg also increases difficulty.