Chapter 6
Regulation

The regulation problem analyzes how quickly a perturbed system returns to its equilibrium setpoint. For this problem, we assume that the setpoint does not change. We can, without loss of generality, assume that the external reference signal is $r = 0$.

With no external reference signal, we can express the general form of the regulation problem as in Fig. 6.1. We take the process, P, as given, subject to uncertainties or disturbances represented by the input, d. We seek an optimal controller, C, with respect to particular design tradeoffs.

6.1 Cost Function

The cost function summarizes the design tradeoffs. We use a cost function based on the $\mathcal{H}_2$ norm, similar to Eq. 5.5. The $\mathcal{H}_2$ norm describes the response of the system to perturbations when averaged over all input frequencies. Minimizing the $\mathcal{H}_2$ norm minimizes the extent to which the system responds to perturbations. Recall that the $\mathcal{H}_2$ norm is often equivalent to the signal energy, which is the total squared deviation of a signal from zero when measured from the time of an initial perturbation until the time when the signal returns to zero.

From Fig. 6.1, the two inputs are the load disturbance, d, and the sensor noise, n. The two outputs are the process output, η, and the control signal, u. We can write the outputs as transfer functions, $\eta(s)$ and $U(s)$, and the cost function in Eq. 5.5 as

$$\mathcal{J} = ||U(s)||_2^2 + \rho^2 ||\eta(s)||_2^2.$$

In this case, we need to relate each of the two outputs to each of the two inputs. We require four transfer functions to describe all of the input–output connections. For the transfer function between the input d and the output η, we write $G_{\eta d}(s)$, for

S. A. Frank, *Control Theory Tutorial*, SpringerBriefs in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-91707-8_6

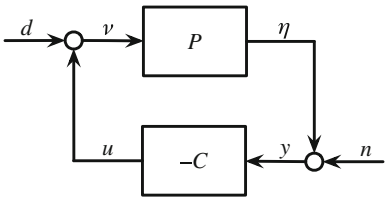

Fig. 6.1 Classic regulation problem illustrated by closed-loop feedback with a constant reference input signal, $r = 0$. The disturbance input, d, perturbs the system process. Such perturbations can be considered as stochasticity in the process, or as uncertainty with regard to the true process dynamics relative to the assumed dynamics. The noise input, n, perturbs the sensor that produces the output measurement, y, based on the actual process output, η. See Fig. 3.2 for context

which we assume that the other input, n, is zero. Using our usual rule for the transfer functions of a closed loop, the four functions are

$$G_{ud} = \frac{-PC}{1+PC} \qquad G_{\eta d} = \frac{P}{1+PC}$$
$$G_{un} = \frac{-C}{1+PC} \qquad G_{\eta n} = \frac{-PC}{1+PC}. \tag{6.1}$$

We can express these transfer functions in terms of the sensitivities in Eq. 3.8 by defining the open loop as $L = PC$, the sensitivity as $S = 1/(1+L)$, and the complementary sensitivity as $T = L/(1+L)$, yielding

$$G_{ud} = -T \qquad G_{\eta d} = PS$$
$$G_{un} = -CS \qquad G_{\eta n} = -T. \tag{6.2}$$

Because $S + T = 1$ at any input, s, these transfer functions highlight the intrinsic design tradeoffs.

We can now consider the total cost as the sum of the response with respect to the input d, holding n at zero, plus the response with respect to the input n, holding d at zero

$$\mathcal{J} = ||G_{ud}(s)||_2^2 + \rho^2 ||G_{\eta d}(s)||_2^2$$
$$+ ||G_{un}(s)||_2^2 + \rho^2 ||G_{\eta n}(s)||_2^2. \tag{6.3}$$

For this example, we use impulse function inputs, $\delta(t)$, which provide a strong instantaneous shock to the system, as defined in the caption of Fig. 4.2. We can design the system to be relatively more or less sensitive to disturbance inputs relative to noise inputs by weighting the disturbance input by μ, so that $d(t) = \mu\delta(t)$ and $n(t) = \delta(t)$. Larger μ causes design by optimization to yield better disturbance regulation at the expense of worse noise regulation.

The transfer function for an impulse is equal to one. Thus, the transfer functions for disturbance and noise inputs are, respectively, $D(s) = \mu$ and $N(s) = 1$. A system's response to an input is simply the product of the input and the system transfer function. For example, the first term in Eq. 6.3 becomes

$$||D(s)G_{ud}(s)||_2^2 = \mu^2||G_{ud}(s)||_2^2,$$

and the full cost function becomes

$$\begin{aligned}\mathcal{J} = \; & \mu^2||G_{ud}(s)||_2^2 + \mu^2\rho^2||G_{\eta d}(s)||_2^2 \\ & + ||G_{un}(s)||_2^2 + \rho^2||G_{\eta n}(s)||_2^2. \end{aligned} \tag{6.4}$$

Using the sensitivity expressions in Eq. 6.2, we can write this expression more simply as

$$\mathcal{J} = ||CS||_2^2 + (\mu^2 + \rho^2)||T||_2^2 + \mu^2\rho^2||PS||_2^2. \tag{6.5}$$

6.2 Optimization Method

This section follows Qiu and Zhou's (2013) optimization algorithm. Their cost function in the final equation on page 31 of their book is equivalent to my cost function in Eq. 6.4.

Optimization finds the controller, $C(s)$, that minimizes the cost function. We search for optimal controllers subject to the constraint that all transfer functions in Eq. 6.1 are stable. Stability requires that the real component be negative for all eigenvalues of each transfer function.

A transfer function's eigenvalues are the roots of the denominator's polynomial in s. For each transfer function in Eq. 6.1, the eigenvalues, s, are obtained by solution of $1 + P(s)C(s) = 0$.

We assume a fixed process, P, and weighting coefficients, μ and ρ. To find the optimal controller, we begin with a general form for the controller, such as

$$C(s) = \frac{q_1 s + q_2}{p_0 s^2 + p_1 s + p_2}. \tag{6.6}$$

We seek the coefficients p and q that minimize the cost function.

Qiu and Zhou (2013) solve the example in which $P(s) = 1/s^2$, for arbitrary values of μ and ρ. The accompanying Mathematica code describes the steps in the solution algorithm. Here, I simply state the solution. Check the article by Qiu and Zhou (2013) and my Mathematica code for the details and for a starting point to apply the optimization algorithm to other problems. The following section applies this method to another example and illustrates the optimized system's response to various inputs.

For $P = 1/s^2$, Qiu and Zhou (2013) give the optimal controller

$$C(s) = \frac{\sqrt{2\rho\mu}\left(\sqrt{\rho} + \sqrt{\mu}\right)s + \rho\mu}{s^2 + \sqrt{2}\left(\sqrt{\rho} + \sqrt{\mu}\right)s + \left(\sqrt{\rho} + \sqrt{\mu}\right)^2},$$

with associated minimized cost,

$$\mathcal{J}^* = \sqrt{2}\left[\mu^2\sqrt{\rho} + \rho^2\sqrt{\mu} + 2\rho\mu\left(\sqrt{\mu} + \sqrt{\rho}\right)\right].$$

For $\rho = 1$, the controller becomes

$$C(s) = \frac{\sqrt{2\mu}\left(1 + \sqrt{\mu}\right)s + \mu}{s^2 + \sqrt{2}\left(1 + \sqrt{\mu}\right)s + \left(1 + \sqrt{\mu}\right)^2}, \tag{6.7}$$

with associated minimized cost,

$$\mathcal{J}^* = \sqrt{2}\left[\mu^2 + \sqrt{\mu} + 2\mu(\sqrt{\mu} + 1)\right].$$

We can see the tradeoffs in design most clearly from the controller with $\rho = 1$. When μ is small, load disturbance inputs are smaller than sensor noise inputs. An optimal system should therefore tolerate greater sensitivity to load disturbances in return for reduced sensitivity to sensor noise.

In the optimal controller described by Eq. 6.7, a small value of μ produces low gain, because $C(s)$ becomes smaller as μ declines. We can see from Eq. 6.1 that a small gain for the controller, C, reduces the sensitivity to noise inputs by lowering G_{un} and $G_{\eta n}$. Similarly, a small gain for C raises the sensitivity of the system output, η, to disturbance inputs by raising $G_{\eta d}$.

The optimal system achieves the prescribed rise in sensitivity to disturbance in order to achieve lower sensitivity to noise.

6.3 Resonance Peak Example

This section applies the previous section's $\mathcal{H}_2$ optimization method to the process

$$P(s) = \frac{1}{s^2 + 0.1s + 1}. \tag{6.8}$$

This process has a resonance peak near $\omega = 1$.

My supplemental Mathematica code derives the optimal controller of the form in Eq. 6.6. The optimal controller is expressed in terms of the cost weightings μ and ρ. The solution has many terms, so there is no benefit in showing it here.

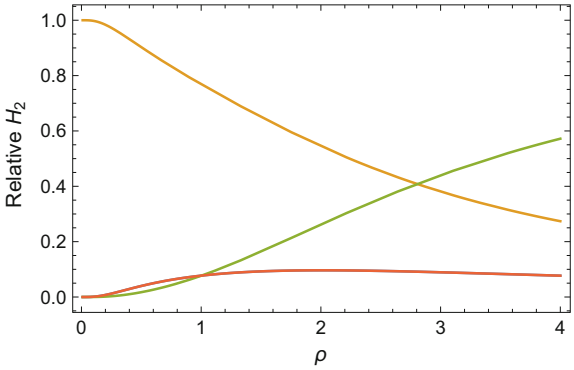

Fig. 6.2 Relative $\mathcal{H}_2$ values for the transfer functions in Eq. 6.1, with $G_{ud} = G_{\eta n}$ in red, $G_{\eta d}$ in gold, and G_{un} in green. The $\mathcal{H}_2$ value for each transfer function is divided by the total $\mathcal{H}_2$ values over all four functions. The transfer functions were derived from the process in Eq. 6.8 and the associated optimal controller. The weighting parameters in the cost function of Eq. 6.4 are $\mu = 1$ and ρ varying along the x-axis of the plot. Swapping values of μ and ρ gives identical results, because of the symmetries in Eqs. 6.1 and 6.4

The general solution in terms of μ and ρ provides a simple way in which to obtain the optimal controller for particular values of μ and ρ. For example, when $\mu = \rho = 1$, the optimal controller is

$$C(s) \approx \frac{0.609(s - 0.81)}{s^2 + 1.73s + 2.49}.$$

Similar controller expressions arise for other values of μ and ρ. Those controllers may be used in the closed loop of Fig. 6.1 to form a complete system.

Figure 6.2 shows the relative $\mathcal{H}_2$ values of the four input–output transfer functions in Eq. 6.1. The $\mathcal{H}_2$ values express sensitivity over all frequencies.

To interpret this figure, look at Eq. 6.4. As the product of the weightings, $\mu\rho$, increases, the output of $G_{\eta d}$ (gold curve) plays an increasingly important role in the total cost relative to the output of G_{un} (green curve).

As the relative cost weighting of $G_{\eta d}$ increases, its $\mathcal{H}_2$ value declines. Similarly, as the relative cost weighting of G_{un} decreases, its $\mathcal{H}_2$ value increases. Once again, we see the sensitivity tradeoffs in response to the relative importance of different perturbations.

The top row of Fig. 6.3 compares the Bode plots for the process, P, and the input–output transfer functions in Eq. 6.1. As ρ increases in the columns from left to right, the rise in the green curve for G_{un} is the strongest change. We can understand that change by examining the cost function in Eq. 6.4. Because $G_{ud} = G_{\eta n}$, a rise in ρ reduces the weighting of G_{un} relative to all other terms.

The strongest increase in relative weighting as ρ rises occurs for $G_{\eta d}$, shown in gold. The mild decline in the gold curve with increasing ρ is consistent with the increased relative cost weighting of that signal.

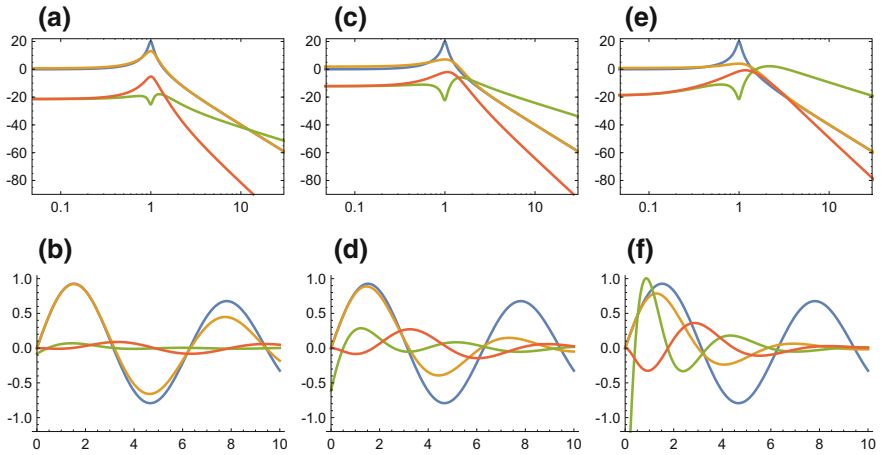

Fig. 6.3 Response of the process in Eq. 6.8 in blue and the transfer functions in Eq. 6.1, with $G_{ud} = G_{\eta n}$ in red, $G_{\eta d}$ in gold, and G_{un} in green. Top row shows Bode magnitude plots. Bottom row shows impulse responses. The input signal weights in Eq. 6.4 are $\mu = 1$ and, for the three columns from left to right, $\rho = 0.25, 1, 4$. Swapping values of μ and ρ gives identical results, because of the symmetries in Eq. 6.1 and 6.4

The bottom row shows the impulse responses. As with the Bode plots, an increase in ρ favors reduced response of $G_{\eta d}$, in gold, causing a smaller impulse response in the right plot with high ρ relative to the left plot with low ρ. Similarly, an increase in ρ weakens the pressure on the G_{un} transfer function in green, causing a larger impulse response with increasing ρ.

6.4 Frequency Weighting

The $\mathcal{H}_2$ norm sums a system's gain over all input frequencies, as in Eq. 5.4. That sum weights all input frequencies equally.

Often, we wish to protect against perturbations that occur primarily in a limited band of frequencies. For example, disturbance loads, d, typically occur at low frequency, reflecting long-term fluctuations or misspecifications in the system's intrinsic processes. In that case, our optimization method should emphasize reducing a system's gain at low frequency with respect to disturbance load inputs and accepting a tradeoff that allows a greater gain at high frequency. By reducing the gain at low frequency, we protect against the common frequencies for load disturbances.

Tradeoffs between low- and high-frequency bands are common. If we start with a process transfer function

$$G(s) = \frac{10(s+1)}{s+10},$$

then at zero frequency, $s = j\omega = 0$, the gain is one. As frequency increases, the gain approaches ten.

If we weight this process transfer function by $W(s) = 1/(s+1)$, then the new system becomes $WG = 10/(s+10)$. Now, the gain declines with increasing frequency, from a maximum of one at zero frequency to a minimum of zero at infinite frequency.

By weighting the original system, G, by the weighting function, W, we cause the $\mathcal{H}_2$ norm of the combined system, WG, to be relatively more sensitive to low-frequency disturbances. When we design a controller by minimizing the $\mathcal{H}_2$ norm associated with WG, we will typically find a system that is better at rejecting low-frequency load disturbances than a design minimizing the $\mathcal{H}_2$ norm associated with G. For the weighted system, optimization will avoid controllers that reject high-frequency load disturbances, because the weighted system already rejects those high-frequency inputs.

Roughly speaking, a weighting function instructs the optimization method to reduce the gain and sensitivity for certain frequencies and to ignore the gain for other frequencies. The weighting functions do not alter the actual system. The weighting functions are only used to alter the cost function and optimization method that determine the optimal controller.

Figure 6.4 shows the regulation feedback system of Fig. 6.1 with additional weightings for the disturbance and noise inputs. The weightings modify the four system transfer functions and associated sensitivities in Eq. 6.2 to be $W_d G_{ud}$, $W_d G_{\eta d}$, $W_n G_{un}$, and $W_n G_{\eta n}$. The cost function in Eq. 6.5 becomes

$$\begin{aligned}\mathcal{J} = \mu^2 ||W_d T||_2^2 + \mu^2 \rho^2 ||W_d P S||_2^2 \\ + ||W_n C S||_2^2 + \rho^2 ||W_n T||_2^2.\end{aligned} \tag{6.9}$$

Consider an example in which we begin with the process, P, in Eq. 6.8. To emphasize low-frequency load disturbances, set $W_d = 1/(s+0.1)$ to be a low-pass filter. That weighting filters out disturbances that are significantly greater than $\omega = 0.1$. To emphasize high-frequency sensor noise, set $W_n = s/(s+10)$. That weighting filters out noise that is significantly less than $\omega = 10$. By using these two filters, the optimization method puts very low weight on any disturbances in midrange frequencies of $\omega = (0.1, 10)$.

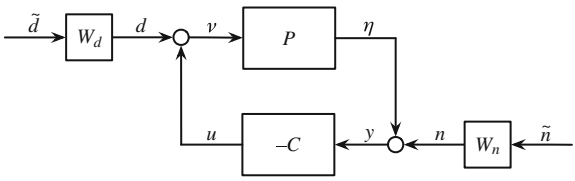

Fig. 6.4 Basic regulation feedback loop in Fig. 6.1 with additional weightings for disturbance and noise inputs. The weightings alter the cost function to emphasize particular frequency bands for disturbance and noise, yielding a modified optimal controller

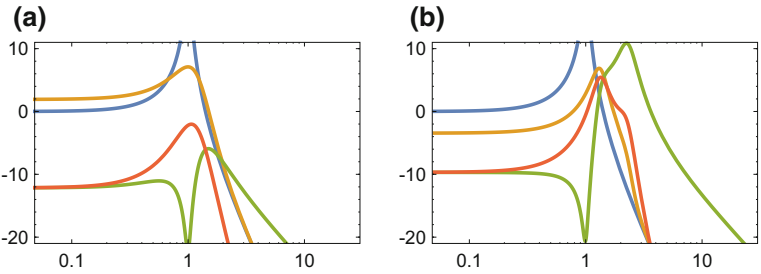

Fig. 6.5 Role of frequency weighted inputs in the design of optimal $\mathcal{H}_2$ controllers for system regulation, illustrated by Bode magnitude plots. **a** Plot for the unweighted case, matching the plot in Fig. 6.3c. **b** Plot for the frequency weighted example in the text, which emphasizes the regulation of low-frequency load disturbances, d, and high-frequency sensor noise, n

By minimizing the weighted $\mathcal{H}_2$ cost in Eq. 6.9, we obtain the optimal controller

$$C(s) = \frac{2.02(s + 1.52)}{s^2 + 1.17s + 6.3}.$$

I calculated the values for this controller by using the numerical minimization function in Mathematica to minimize the $\mathcal{H}_2$ cost, subject to the constraint that all transfer functions in Eq. 6.1 are stable. See the supplemental Mathematica code.

Figure 6.5 compares the optimized system response for the unweighted and weighted cases. Panel **a** shows the Bode magnitude response of the optimized system for the unweighted case, equivalent to the plot in Fig. 6.3c. Panel **b** shows the response of the optimized system for the weighted case in this section.

The weighted case emphasizes low-frequency load disturbances and high-frequency sensor noise, with low weight on midrange frequencies. Comparing the unweighted case in (a) with the weighted case in (b), we see two key differences.

First, the weighted case allows a large rise in magnitudes and associated sensitivity to perturbations for midrange frequencies. That rise occurs because the particular weighting functions in this example discount midrange perturbations.

Second, the gold curve shows that the weighted case significantly reduces the low-frequency sensitivity of system outputs, η, to load disturbances, d. The gold curve describes the response of the transfer function, $G_{\eta d}$. Note that, because of the log scaling for magnitude, almost all of the costs arise in the upper part of the plot. The low relative magnitude for the lower part contributes little to the overall cost.