Chapter 5
Performance and Robustness Measures

A theory of design tradeoffs requires broadly applicable measures of cost, performance, stability, and robustness. For example, the PID controller in the previous example performs reasonably well, but we ignored costs. That PID controller achieved good tracking performance by using high gain amplification of low-frequency input signals. High gain in a negative feedback loop quickly drives the error to zero.

High gain has two potential problems. First, high signal amplification may require excessive energy in physical or biological systems. We must consider those costs for a high gain controller.

Second, high gain can cause system instability, with potential for system failure. We must consider the tradeoff between the benefits of high gain and the loss of robustness against perturbations or uncertainties in system dynamics.

Beyond the simple PID example, we must consider a variety of tradeoffs in performance and robustness (Zhou and Doyle 1998; Qiu and Zhou 2010). Earlier, I discussed tradeoffs in system sensitivities to disturbance and noise. I also presented qualitative descriptions of system performance in terms of response time and tracking performance.

To advance the theory, we need specific measures of cost, performance, stability and robustness. We also need techniques to find optimal designs in relation to those conflicting measures of system attributes.

We will never find a perfect universal approach. There are too many dimensions of costs and benefits, and too many alternative ways to measure system attributes. Nonetheless, basic measures and simple optimization methods provide considerable insight into the nature of design. Those insights apply both to the building of human-designed systems to achieve engineering goals and to the interpretation and understanding of naturally designed biological systems built by evolutionary processes.

© The Author(s) 2018

S. A. Frank, *Control Theory Tutorial*, SpringerBriefs in Applied Sciences and Technology, https://doi.org/10.1007/978-3-319-91707-8_5

5.1 Performance and Cost: $\mathcal{J}$

To analyze performance, we must measure the costs and benefits associated with a particular system. We often measure those costs and benefits by the distance between a system's trajectory and some idealized trajectory with zero cost and perfect performance.

Squared deviations provide a distance measure between the actual trajectory and the idealized trajectory. Consider, for example, the control signal, $u(t)$, which the controller produces to feed into the system process, as in Fig. 2.1c.

The value of $|u(t)|^2 = u^2$ measures the magnitude of the signal as a squared distance from zero. We can think of u^2 as the instantaneous power of the control signal. Typically, the power requirements for control are a cost to be minimized.

The square of the error output signal, $|e(t)|^2 = e^2$, measures the distance of the system from the ideal performance of $e = 0$. Minimizing the squared error maximizes performance. Thus, we may think of performance at any particular instant, t, in terms of the cost function

$$\mathcal{J}(t) = u^2 + \rho^2 e^2,$$

for which minimum cost corresponds to maximum performance. Here, ρ is a weighting factor that determines the relative value of minimizing the control signal power, u^2, versus minimizing the tracking error, e^2.

Typically, we measure the cost function over a time interval. Summing up $\mathcal{J}(t)$ continuously from $t = 0$ to T yields

$$\mathcal{J} = \int_0^T (u^2 + \rho^2 e^2)\mathrm{d}t. \tag{5.1}$$

Most squared distance or quadratic performance analyses arise from extensions of this basic equation. Given this measure, optimal design trades off minimizing the energy cost to drive the system versus maximizing the benefit of tracking a target goal.

5.2 Performance Metrics: Energy and $\mathcal{H}_2$

The cost measure in Eq. 5.1 analyzes signals with respect to time. It is natural to think of inputs and outputs as changing over time. With temporal dynamics, we can easily incorporate multivariate signals and nonlinearities. In spite of those advantages, we often obtain greater insight by switching to a frequency analysis of signals, as in the previous chapters.

In this section, I present alternative measures of cost and performance in terms of transfer functions and complex signals. Those alternative measures emphasize frequencies of fluctuations rather than changes through time. Frequency and complex analysis allow us to take advantage of transfer functions, Bode plots, and other powerful analytical tools that arise when we assume linear dynamics.

The assumption of linearity does not mean that we think the actual dynamics of physical and biological processes are linear. Instead, starting with the linear case provides a powerful way in which to gain insight about dynamics.

In the previous section, we considered how to measure the magnitude of fluctuating control and error signals. A magnitude that summarizes some key measure is often called a *norm*. In the prior section, we chose the sum of squared deviations from zero, which is related to the 2–norm of a signal

$$\|u(t)\|_2 = \left(\int_0^\infty |u(t)|^2 \mathrm{d}t\right)^{1/2}. \tag{5.2}$$

The energy of the signal is the square of the 2–norm, $\|u(t)\|_2^2$. When the time period in the cost function of Eq. 5.1 goes to infinity, $T \to \infty$, we can write the cost function as

$$\mathcal{J} = \|u(t)\|_2^2 + \rho^2 \|e(t)\|_2^2. \tag{5.3}$$

The signal $u(t)$ is a function of time. The associated transfer function $U(s)$ describes exactly the same signal, but as a function of the complex number, s, rather than of time, t.

It is often much easier to work with the transfer function for analysis, noting that we can go back and forth between time and transfer function descriptions. For the analysis of squared distance metrics, the 2–norm of the transfer function expression is

$$\|U(s)\|_2 = \left(\frac{1}{2\pi}\int_{-\infty}^\infty |U(j\omega)|^2 \mathrm{d}\omega\right)^{1/2}. \tag{5.4}$$

This transfer function 2–norm is often referred to as the $\mathcal{H}_2$ norm. The term $|U(j\omega)|^2$ is the square of the Bode gain or magnitude, as in Fig. 2.2e. That gain describes the amplification of a sinusoidal input at frequency ω. The $\mathcal{H}_2$ norm expresses the average amplification of input signals over all input frequencies.

If the goal is to minimize the control input signal, u, or the error deviation from zero, e, then the greater the amplification of a signal, the greater the cost. Thus, we can use the $\mathcal{H}_2$ norm to define an alternative cost function as

$$\mathcal{J} = \|U(s)\|_2^2 + \rho^2 \|E(s)\|_2^2, \tag{5.5}$$

which leads to methods that are often called $\mathcal{H}_2$ analysis. This cost describes the amplification of input signals with respect to control and error outputs when averaged overall input frequencies. Minimizing this cost reduces the average amplification of input signals.

If the energy 2–norm in Eq. 5.2 is finite, then the energy 2–norm and the $\mathcal{H}_2$ norm are equivalent, $\|u(t)\|_2 = \|U(s)\|_2$, and we can use Eqs. 5.3 and 5.5 interchangeably. Often, it is more convenient to work with the transfer function form of the $\mathcal{H}_2$ norm.

We can use any combination of signals in the cost functions. And we can use different weightings for the relative importance of various signals. Thus, the cost functions provide a method to analyze a variety of tradeoffs.

5.3 Technical Aspects of Energy and $\mathcal{H}_2$ Norms

I have given three different cost functions. The first in Eq 5.1 analyzes temporal changes in signals, such as $u(t)$, over a finite time interval. That cost function is the most general, in the sense that we can apply it to any finite signals. We do not require assumptions about linearity or other special attributes of the processes that create the signals.

The second function in Eq. 5.3 measures cost over an infinite time interval and is otherwise identical to the first measure. Why consider the unrealistic case of infinite time?

Often, analysis focuses on a perturbation that moves a stable system away from its equilibrium state. As the system returns to equilibrium, the error and control signals go to zero. Thus, the signals have positive magnitude only over a finite time period, and the signal energy remains finite. As noted above, if the energy 2–norm is finite, then the energy 2–norm and the $\mathcal{H}_2$ norm are equivalent, and the third cost function in Eq. 5.5 is equivalent to the second cost function in Eq. 5.3.

If the signal energy of the second cost function in Eq. 5.3 is infinite, then that cost function is not useful. In an unstable system, the error often grows with time, leading to infinite energy of the error signal. For example, the transfer function $1/(s-1)$ has temporal dynamics given by $y(t) = y(0)e^t$, growing exponentially with time. The system continuously amplifies an input signal, creating instability and an output signal with infinite energy.

When the energy is infinite, the $\mathcal{H}_2$ norm may remain finite. For the transfer function $1/(s-1)$, the $\mathcal{H}_2$ norm is $1/\sqrt{2}$. The average amplification of signals remains finite. In general, for a transfer function, $G(s)$, the $\mathcal{H}_2$ norm remains finite as long as $G(j\omega)$ does not go to infinity for any value of ω, and $G(j\omega) \to 0$ as $\omega \to \pm\infty$. Thus, the $\mathcal{H}_2$ norm cost in Eq. 5.5 can be used in a wider range of applications.

The $\mathcal{H}_2$ norm is related to many common aspects of signal processing and time series analysis, such as Fourier analysis, spectral density, and autocorrelation.

5.4 Robustness and Stability: $\mathcal{H}_\infty$

A transfer function for a system, $G(s)$, defines the system's amplification of input signals. For a sinusoidal input at frequency ω, the amplification, or gain, is the absolute value of the transfer function at that frequency, $|G(j\omega)|$.

Often, the smaller a system's amplification of inputs, the more robust the system is against perturbations. Thus, one common optimization method for designing controllers seeks to minimize a system's greatest amplification of inputs. Minimizing the greatest amplification guarantees a certain level of protection against the worst case perturbation. In some situations, one can also guarantee that a system is stable if its maximum signal amplification is held below a key threshold.

A system's maximum amplification of sinusoidal inputs over all input frequencies, ω, is called its $\mathcal{H}_\infty$ norm. For a system $G(s)$, the $\mathcal{H}_\infty$ norm is written as $\|G(s)\|_\infty$. The norm describes the maximum of $|G(j\omega)|$ over all ω. The maximum is also the peak gain on a Bode magnitude plot, which is equivalent to the resonance peak.

System stability and protection against perturbations set two fundamental criteria for system design. Thus, $\mathcal{H}_\infty$ methods are widely used in the engineering design of controllers and system architectures (Zhou and Doyle 1998).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.