Fundamentals of Linear Control

A Concise Approach

MAURÍCIO C. DE OLIVEIRA
University of California, San Diego

CAMBRIDGE
UNIVERSITY PRESS

1 Introduction

In controls we make use of the abstract concept of a *system*: we identify a phenomenon or a process, the *system*, and two classes of *signals*, which we label as *inputs* and *outputs*. A signal is something that can be measured or quantified. In this book we use real numbers to quantify signals. The classification of a particular signal as an input means that it can be identified as the *cause* of a particular system behavior, whereas an output signal is seen as the *product* or *consequence* of the behavior. Of course the classification of a phenomenon as a system and the labeling of input and output signals is an abstract construction. A mathematical description of a system and its signals is what constitutes a *model*. The entire abstract construction, and not only the equations that we will later associate with particular signals and systems, is the model.

We often represent the relationship between a system and its input and output signals in the form of a *block-diagram*, such as the ones in Fig. 1.1 through Fig. 1.3. The diagram in Fig. 1.1 indicates that a system, G, produces an output signal, y, in the presence of the input signal, u. Block-diagrams will be used to represent the interconnection of systems and even algorithms. For example, Fig. 1.2 depicts the components and signals in a familiar controlled system, a water heater; the block-diagram in Fig. 1.3 depicts an algorithm for converting temperature in degrees Fahrenheit to degrees Celsius, in which the output of the circle in Fig. 1.3 is the algebraic sum of the incoming signals with signs as indicated near the incoming arrows.

1.1 Models and Experiments

Systems, signals, and models are often associated with concrete or abstract experiments. A model reflects a particular setup in which the outputs appear *correlated* with a prescribed set of inputs. For example, we might attempt to model a car by performing the following experiment: on an unobstructed and level road, we depress the accelerator pedal and let the car travel in a straight line.[1] We keep the pedal excursion constant and let the car reach constant velocity. We record the amount the pedal has been depressed and the car's terminal velocity. The results of this experiment, repeated multiple times with different amounts of pedal excursion, might look like the data shown in Fig. 1.4. In this experiment the signals are

[1] This may bring to memory a bad joke about physicists and spherical cows . . .

Figure 1.1 System represented as a block-diagram; u is the input signal; y is the output signal; y and u are related through $y = G(u)$ or simply $y = Gu$.

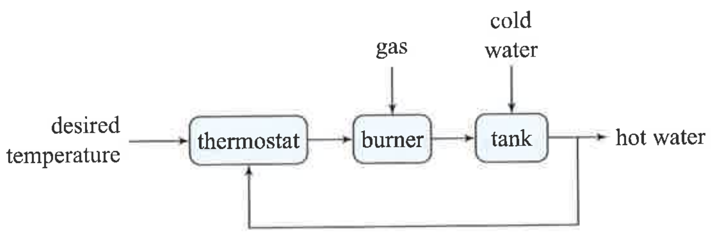


Figure 1.2 Block-diagram of a controlled system: a gas water heater; the blocks thermostat, burner, and tank, represent components or sub-systems; the arrows represent the *flow* of input and output signals.

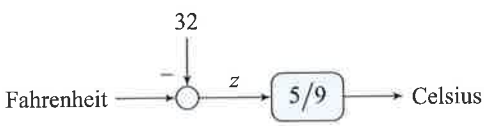


Figure 1.3 Block-diagram of an algorithm to convert temperatures in Fahrenheit to Celsius: Celsius $= 5/9(\text{Fahrenheit} - 32)$; the output of the circle block is the algebraic sum of the incoming signals with the indicated sign, i.e. $z = \text{Fahrenheit} - 32$.

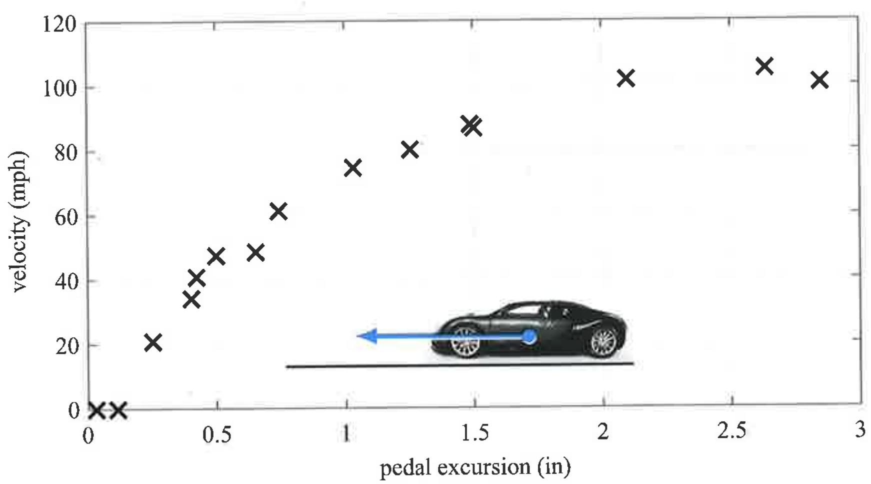


Figure 1.4 Experimental determination of the effect of pressing the gas pedal on the car's terminal velocity; the pedal excursion is the input signal, u, and the car's terminal velocity is the output signal, y.

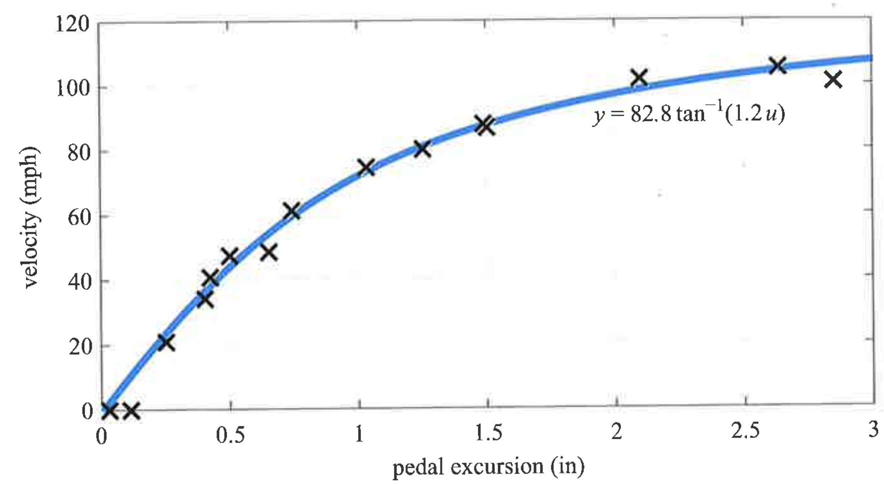


Figure 1.5 Fitting the curve $y = \alpha \tan^{-1}(\beta u)$ to the data from Fig. 1.4.

input: pedal excursion, in cm, inches, etc.;
output: terminal velocity of the car, in m/s, mph, etc.

The **system** is the car *and* the particular conditions of the experiment. The data captures the fact that the car does not move at all for small pedal excursions and that the terminal velocity *saturates* as the pedal reaches the end of its excursion range.

From Fig. 1.4, one might try to *fit* a particular mathematical function to the experimental data[2] in hope of obtaining a *mathematical model*. In doing so, one invariably loses something in the name of a simpler description. Such trade-offs are commonplace in science, and it should be no different in the analysis and design of control systems. Figure 1.5 shows the result of fitting a curve of the form

$$y = \alpha \tan^{-1}(\beta u),$$

where u is the input, pedal excursion in inches, and y is the output, terminal velocity in mph. The parameters $\alpha = 82.8$ and $\beta = 1.2$ shown in Fig. 1.5 were obtained from a standard least-squares fit. See also P1.11.

The choice of the above particular function involving the arc-tangent might seem somewhat arbitrary. When possible, one should select candidate functions from first principles derived from physics or other scientific reasoning, but this does not seem to be easy to do in the case of the experiment we described. Detailed physical modeling of the vehicle would involve knowledge and further modeling of the components of the vehicle, not to mention the many uncertainties brought in by the environment, such as wind, road conditions, temperature, etc. Instead, we make an "educated choice" based on certain physical aspects of the experiment that we believe the model should capture. In this case, from our daily experience with vehicles, we expect that the terminal velocity

[2] All data used to produce the figures in this book is available for download from the website http://www.cambridge.org/deOliveira.

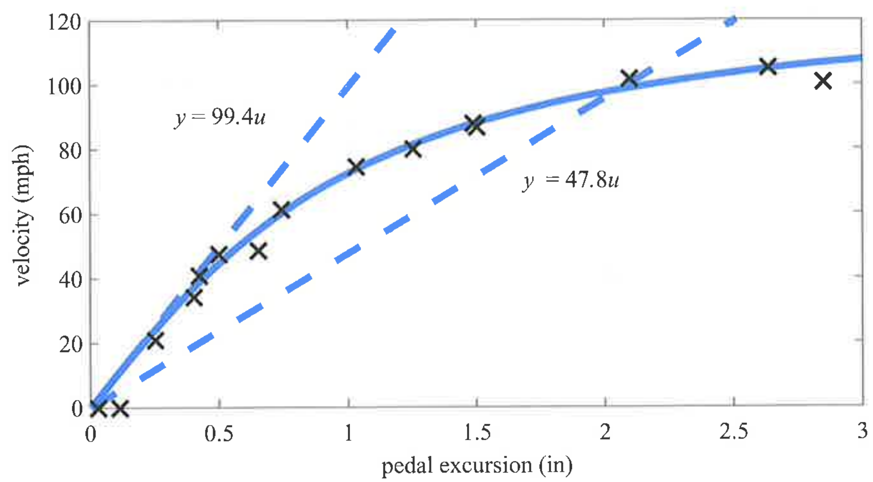


Figure 1.6 Linear mathematical models of the form $y = \gamma u$ for the data in Fig. 1.4 (dashed); the model with $\gamma = 47.8$ was obtained by a least-squares fit; the model with $\gamma = 99.4$ was obtained after linearization of the nonlinear model (solid) obtained in Fig. 1.5; see P1.12 and P1.11.

will eventually *saturate*, either as one reaches full throttle or as a result of limitations on the maximum power that can be delivered by the vehicle's powertrain. We also expect that the function be *monotone*, that is, the more you press the pedal, the larger the terminal velocity will be. Our previous exposure to the properties of the arc-tangent function *and* engineering intuition about the expected outcome of the experiment allowed us to successfully select this function as a suitable candidate for a model.

Other families of functions might suit the data in Fig. 1.5. For example, we could have used *polynomials*, perhaps constrained to pass through the origin and ensure monotonicity. One of the most useful classes of mathematical models one can consider is that of *linear models*, which are, of course, first-order polynomials. One might be tempted to equate linear with simple. Whether or not this might be true in some cases, simplicity is far from a sin. More often than not, the loss of some feature neglected by a linear model is offset by the availability of a much broader set of analytic tools. It is better to *know* when you are wrong than to *believe* you are right. As the title suggests, this book is mostly concerned with linear models. Speaking of linear models, one might propose describing the data in Fig. 1.4 by a linear mathematical model of the form

$$y = \gamma u. \tag{1.1}$$

Figure 1.6 shows two such models (dashed lines). The curve with slope coefficient $\gamma = 47.8$ was obtained by performing a least-squares fit to all data points (see P1.11). The curve with coefficient $\gamma = 99.4$ is a first-order approximation of the nonlinear model calculated in Fig. 1.5 (see P1.12). Clearly, each model has its limitations in describing the experiment. Moreover, one model might be better suited to describe certain aspects of the experiment than the other. Responsibility rests with the engineer or the scientist to select the model, or perhaps set of models, that better fits the problem in hand, a task that at times may resemble an art more than a science.

1.2 Cautionary Note

It goes without saying that the mathematical models described in Section 1.1 do not purport to capture every detail of the experiment, not to mention reality. Good models are the ones that capture *essential* aspects that we *perceive* or can experimentally validate as real, for example how the terminal velocity of a car responds to the acceleration pedal in the given experimental conditions. A model does not even need to be *correct* to be useful: for centuries humans used[3] a model in which the sun revolves around the earth to predict and control their days! What is important is that models provide a way to express *relevant* aspects of reality using mathematics. When mathematical models are used in control design, it is therefore with the understanding that the model is bound to capture only a subset of features of the actual phenomenon they represent. At no time should one be fooled into *believing* in a model. The curious reader will appreciate [Fey86] and the amusingly provocative [Tal07].

With this caveat in mind, it is useful to think of an idealized *true* or *nominal model*, just as is done in physics, against which a particular setup can be *mathematically* evaluated. This nominal model might even be different than the model used by a particular control algorithm, for instance, having more details or being more complex or more accurate. Of course *physical* evaluation of a control system with respect to the underlying natural phenomenon is possible only by means of experimentation which should also include the physical realization of the controller in the form of computer hardware and software, electric circuits, and other necessary mechanical devices. We will discuss in Chapter 5 how certain physical devices can be used to implement the dynamic controllers you will learn to design in this book.

The models discussed so far have been *static*, meaning that the relationship between inputs and outputs is *instantaneous* and is independent of the past history of the system or their signals. Yet the main objective of this book is to work with *dynamic* models, in which the relationship between present inputs and outputs may depend on the present and past history[4] of the signals.

With the goal of introducing the main ideas behind feedback control in a simpler setup, we will continue to work with static models for the remainder of this chapter. In the case of static models, a mathematical *function* or a set of *algebraic equations* will be used to represent such relationships, as done in the models discussed just above in Section 1.1.

Dynamic models will be considered starting in Chapter 2. In this book, signals will be continuous functions of time, and dynamic models will be formulated with the help of *ordinary differential equations*. As one might expect, experimental procedures that can estimate the parameters of dynamic systems need to be much more sophisticated than the ones discussed so far. A simple experimental procedure will be briefly discussed in Section 2.4, but the interested reader is encouraged to consult one of the many excellent works on this subject, e.g. [Lju99].

[3] Apparently 1 in 4 Americans and 1 in 3 Europeans still go by that model [Gro14].
[4] What about the future?

1.3 A Control Problem

Consider the following problem:

Under the experimental conditions described in Section 1.1 and given a target terminal velocity, $\bar{y}$, is it possible to design a system, the controller, that is able to command the accelerator pedal of a car, the input, u, to produce a terminal velocity, the output, y, equal to the target velocity?

An *automatic* system that can solve this problem is found in many modern cars, with the name *cruise controller*. Of course, another *system* that is capable of solving the same problem is a human driver.[5] In this book we are mostly interested in solutions that can be implemented as an *automatic control*, that is, which can be performed by some combination of mechanical, electric, hydraulic, or pneumatic systems running without human intervention, often being programmed in a digital computer or some other logical circuit or calculator.

Problems such as this are referred to in the control literature as *tracking* problems: the controller should make the system, a car, *follow* or *track* a given target output, the desired terminal velocity. In the next sections we will discuss two possible approaches to the cruise control problem.

1.4 Solution without Feedback

The role of the controller in tracking is to compute the input signal u which produces the desired output signal y. One might therefore attempt to solve a tracking problem using a system (controller) of the form

$$u = K(\bar{y}).$$

This controller can use only the reference signal, the target output $\bar{y}$, and is said to be in *open-loop*,[6] as the controller output signal, u, is not a function of the system output signal, y.

With the intent of analyzing the proposed solution using mathematical models, assume that the car can be represented by a *nominal model*, say G, that relates the input u (pedal excursion) to the output y (terminal velocity) through the mathematical function

$$y = G(u).$$

The connection of the controller with this idealized model is depicted in the block-diagram in Fig. 1.7. Here the function G can be obtained after fitting experimental data as done in Figs. 1.5 and 1.6, or borrowed from physics or engineering science principles.

[5] After some 16 years of *learning*.

[6] As opposed to *closed-loop*, which will be discussed in Section 1.5.

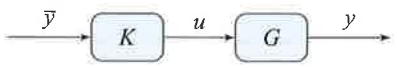


Figure 1.7 Open-loop control: the controller, K, is a function of the reference input, $\bar{y}$, but not a function of the system output, y.

The block-diagram in Fig. 1.7 represents the following relationships:

$$y = G(u), \qquad u = K(\bar{y}),$$

that can be combined to obtain

$$y = G\left(K(\bar{y})\right).$$

If G is *invertible* and K is chosen to be the inverse of G, that is $K = G^{-1}$, then

$$y = G\left(G^{-1}(\bar{y})\right) = \bar{y}.$$

Matching the controller, K, with the nominal model, G, is paramount: if $K \neq G^{-1}$ then $y \neq \bar{y}$.

When both the nominal model G and the controller K are linear,

$$y = G\,u, \qquad u = K\,\bar{y}, \qquad y = G\,K\,\bar{y},$$

from which $\bar{y} = y$ only if the product of the *constants* K and G is equal to one:

$$K\,G = 1 \qquad \Longrightarrow \qquad K = G^{-1}, \qquad u = G^{-1}\,\bar{y}.$$

Because the control law relies on knowledge of the nominal model G to achieve its goal, any imperfection in the model or in the implementation of the controller will lead to less than perfect tracking.

1.5 Solution with Feedback

The controller in the open-loop solution considered in Section 1.4 is allowed to make use only of the target output, $\bar{y}$. When a measurement, even if imprecise, of the system output is available, one may benefit from allowing the controller to make use of the measurement signal, y. In the case of the car cruise control, the terminal velocity, y, can be measured by an on-board speedometer. Of course the target velocity, $\bar{y}$, is set by the driver.

Controllers that make use of output signals to compute the control inputs are called *feedback controllers*. In its most general form, a feedback controller has the functional form

$$u = K(\bar{y}, y).$$

In practice, most feedback controllers work by first creating an *error signal*, $\bar{y} - y$, which is then used by the controller:

$$u = K(e), \qquad e = \bar{y} - y. \tag{1.2}$$

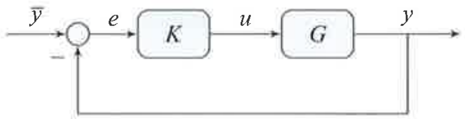


Figure 1.8 Closed-loop feedback control: the controller, K, is a function of the reference input, $\bar{y}$, and the system output, y, by way of the error signal, $e = \bar{y} - y$.

This scheme is depicted in the block-diagram in Fig. 1.8. One should question whether it is possible to implement a physical system that replicates the block-diagram in Fig. 1.8. In this diagram, the measurement, y, that takes part in the computation of the control, u, in the controller block, K, is the same as that which comes out of the system, G. In other words, the signals flow in this diagram is *instantaneous*. Even though we are not yet properly equipped to address this question, we anticipate that it will be possible to construct and analyze *implementable* or *realizable* versions of the feedback diagram in Fig. 1.8 by taking into account dynamic phenomena, which we will start discussing in the next chapter.

At this point, we are content to say that if the computation implied by feedback is performed *fast enough*, then the scheme *should* work. We analyze the proposed feedback solution only in the case of static linear models, that is, when both the controller, K, and the system to be controlled, G, are linear. Feedback controllers of the form (1.2), which are *linear* and *static*, are known by the name *proportional controllers*, or P controllers for short. In the *closed-loop* diagram of Fig. 1.8, we can think of the signal $\bar{y}$, the target velocity, as an input, and of the signal y, the terminal velocity, as an output. A mathematical description of the relationship between the input signal, $\bar{y}$, and output signal, y, assuming linear models, can be computed from the diagram:

$$y = Gu, \qquad u = Ke, \qquad e = \bar{y} - y.$$

After eliminating the signals e and u we obtain

$$y = GKe = GK(\bar{y} - y) \qquad \Longrightarrow \qquad (1 + GK)y = GK\bar{y}.$$

When $GK \neq -1$,

$$y = H\bar{y}, \qquad H = \frac{GK}{1 + GK}.$$

A mathematical relationship governing a particular pair of inputs and outputs is called a *transfer-function*. The function H calculated above is known as a *closed-loop transfer-function*.

Ironically, a first conclusion from the closed-loop analysis is that it is not possible to achieve exact tracking of the target velocity since H cannot be equal to one for any finite value of the constants G and K, not even when $K = G^{-1}$, which was the open-loop solution. However, it is not so hard to make H get close to one: just make K large! More precisely, make the product GK large. How large it needs to be depends on the particular system G. However, a welcome side-effect of the closed-loop solution is that the controller gain, K, does not depend directly on the value of the system model, G.

Table 1.1 Closed-loop transfer-function, H, for various values of K and G

	K				
G	0.02	0.05	0.5	1	3
47.8	0.4888	0.7050	0.9598	0.9795	0.9931
73.3	0.5945	0.7856	0.9734	0.9865	0.9955
99.4	0.6653	0.8325	0.9803	0.9900	0.9967

As the calculations in Table 1.1 reveal, the closed-loop transfer-function, H, remains within 1% of 1 for values K greater than or equal to 3 for *any* value of G lying between the two crude linear models estimated earlier in Fig. 1.6.

In other words, feedback control does not seem to rely on exact knowledge of the system model in order to achieve good tracking performance. This is a major feature of feedback control, and one of the reasons why we may get away with using incomplete and not extremely accurate mathematical models for feedback design. One might find this strange, especially to scientists and engineers trained to look for accuracy and fidelity in their models of the world, a line of thought that might lead one to believe that better accuracy *requires* the use of complex models. For example, the complexity required for accurately modeling the interaction of an aircraft with its surrounding air may be phenomenal. Yet, as the Wright brothers and other flight pioneers demonstrated, it is possible to design and implement effective feedback control of aircraft without relying explicitly on such complex models.

This remarkable feature remains for the most part true even if nonlinear[7] models are considered, although the computation of the transfer-function, H, becomes more complicated.[8] Figure 1.9 shows a plot of the ratio $y/\bar{y}$ for various choices of gain, K, when a linear controller is in feedback with the static nonlinear model, G, fitted in Fig. 1.5. The trends are virtually the same as those obtained using linear models. Note also that the values of the ratio of the terminal velocity by the target velocity are close to the values of H calculated for the linear model with gain $G = 99.4$ which was obtained through "linearization" of the nonlinear model, especially at low velocities.

Insight on the reasons why feedback control can achieve tracking without relying on precise models is obtained if we look at the control, the signal u, that is effectively computed by the closed-loop solution. Following steps similar to the ones used in the derivation of the closed-loop transfer-function, we calculate

$$u = Ke = K(\bar{y} - y) = K(1 - H)\bar{y} = \frac{K}{1 + GK}\bar{y} = \frac{1}{K^{-1} + G}\bar{y}.$$

Note that $\lim_{K \to \infty} u = G^{-1}\bar{y}$, which is exactly the same control as that computed in open-loop (see Section 1.4). This time, however, it is the feedback loop that *computes* the function G^{-1} based on the error signal, $e = \bar{y} - y$. Indeed, u is simply equal to

[7] Many but not all nonlinear models.

[8] It requires solving the nonlinear algebraic equation $y = G(K(\bar{y} - y))$ for y. The dynamic version of this problem is significantly more complex.

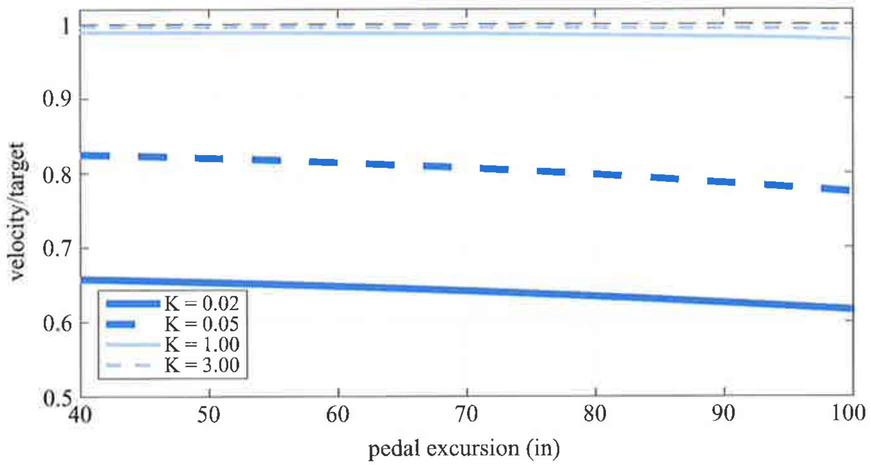


Figure 1.9 Effect of the gain K on the ability of the terminal velocity, y, to track a given target velocity, $\bar{y}$, when the linear feedback control, $u = K(y - \bar{y})$, is in closed-loop (Fig. 1.8) with the nonlinear model, $y = G(u) = 82.8 \tan^{-1}(1.2u)$ from Fig. 1.5.

$K(\bar{y} - y)$, which, when K is made large, converges to $G^{-1}\bar{y}$ by virtue of feedback, no matter what the value of G is. A natural question is what are the side-effects of raising the control gain in order to improve the tracking performance? We will come back to this question at many points in this book as we learn more about dynamic systems and feedback.

1.6 Sensitivity

In previous sections, we made statements regarding how insensitive the closed-loop feedback solution was with respect to changes in the system model when compared with the open-loop solution. We can quantify this statement in the case of static linear models.

As seen before, in both open- and closed-loop solutions to the tracking control problem, the output y is related to the target output $\bar{y}$ through

$$y = H(G)\bar{y}.$$

The notation $H(G)$ indicates that the transfer-function, H, depends on the system model, G. In the open-loop solution $H(G) = GK$ and in the closed-loop solution $H(G) = GK(1 + GK)^{-1}$.

Now consider that G assumes values in the neighborhood of a certain nominal model $\bar{G}$ and that $H(\bar{G}) \neq 0$. Assume that those changes in G affect H in a continuous and differentiable way so that[9]

$$H(G) = H(\bar{G}) + H'(\bar{G})(\Delta G) + O\left(\Delta G^2\right), \qquad \Delta G = G - \bar{G},$$

[9] The notation $O(x^n)$ indicates a polynomial in x that has only terms with degree greater than or equal to n.

which is the Taylor series expansion of H as a function of G about $\bar{G}$. Discarding terms of order greater than or equal to two, we can write

$$H(\bar{G}) - H(G) \approx H'(\bar{G})(\bar{G} - G).$$

After dividing by $H(\bar{G})$ we obtain an expression for how changes in G affect the transfer-function $H(G)$:

$$\frac{H(\bar{G}) - H(G)}{H(\bar{G})} \approx S(\bar{G})\frac{\bar{G} - G}{\bar{G}}, \qquad S(G) = \frac{G}{H(G)}H'(G).$$

The function S is called the *sensitivity* function.

Using this formula we compute the sensitivity of the open-loop solution. In the case of linear models,

$$H(G) = GK, \qquad \Longrightarrow \qquad S(G) = \frac{G}{GK}K = 1.$$

This can be interpreted as follows: in open-loop, a relative change in the system model, G, produces a relative change in the output, y, of the same order.

In closed-loop, after some calculations (see P1.13),

$$H(G) = \frac{GK}{1 + GK}, \qquad \Longrightarrow \qquad S(G) = \frac{1}{1 + GK}. \tag{1.3}$$

By making K large we not only improve the tracking performance but also reduce the sensitivity S. Note that $S + H = 1$, hence $S = 1 - H$, so that the values of S can be easily calculated from Table 1.1 in the case of the car cruise control. For this reason, H is known as the *complementary sensitivity* function.

In the closed-loop diagram of Fig. 1.8, the transfer-function from the reference input, $\bar{y}$, to the tracking error, e, is

$$e = \bar{y} - y = (1 - H)\bar{y} = S\bar{y},$$

which is precisely the sensitivity transfer-function that we have calculated based on model variations. The smaller the sensitivity, S, the better the controller tracks the reference input, $\bar{y}$. Perfect tracking would be achieved if we could make $S = 0$. This is a win–win *coincidence*: the closer controllers can track references, the less sensitive the closed-loop will be to variations in the model.

1.7 Disturbances

Another way of accounting for variability in models is to introduce additional *disturbance signals*. Consider, for example, the block-diagram in Fig. 1.10, in which the disturbance signal, w, adds to the input signal, u, that is,

$$y = G(u + w).$$

It is the scientist or engineer who must distinguish disturbance signals from regular input signals. Disturbances are usually nuisances that might be present during the operation of the system but are hard to model at the design phase, as well as other phenomena

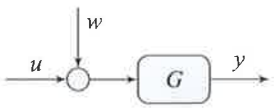


Figure 1.10 System with input disturbance w.

affecting the system that are not completely understood. For example, in Section 2.7, we will use a disturbance signal entering the block-diagram in Fig. 1.10 as w to model a road slope in the car cruise control problem. Because feedback control can be very effective in handling disturbances, delegating difficult aspects of a problem to disturbances is key to simplifying the control design process. Indeed, excerpts from the 1903 Wright brothers' patent for a *flying machine*, shown in Fig. 1.11, hint that this way of thinking might have played a central role in the conquest of flight.

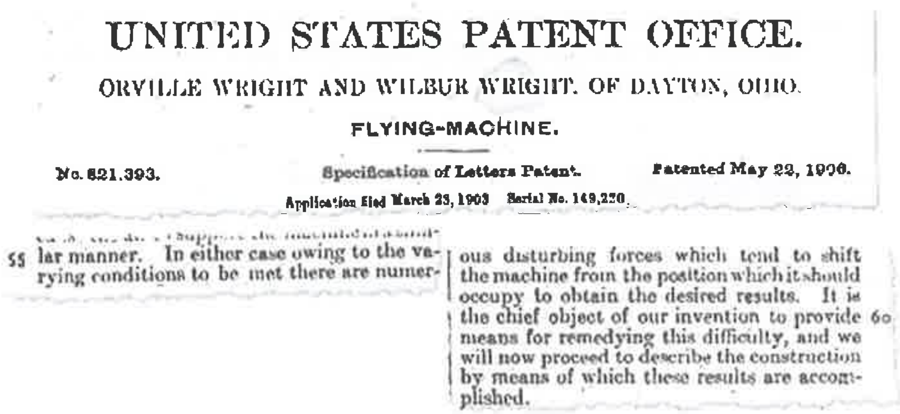
It is easy to incorporate disturbances into the basic open- and closed-loop schemes of Figs. 1.7 and 1.8, which we do in Figs. 1.12 and 1.13. In both cases, one can write the output, y, in terms of the reference input, $\bar{y}$, and the disturbance, w. Better yet, we can write the transfer-function from the inputs, $\bar{y}$ and w, to the tracking error, $e = \bar{y} - y$. In open-loop we calculate with Fig. 1.12 that

$$e = \bar{y} - y = \bar{y} - G(K\bar{y} + w) = (1 - GK)\bar{y} - Gw.$$

Substituting the proposed open-loop solution, $K = G^{-1}$, we obtain

$$e = -Gw,$$

which means that open-loop control is very effective at tracking but has no capability to *reject* the disturbance w, as one could have anticipated from the block-diagram in Fig. 1.12. Open-loop controllers will perform poorly in the presence of disturbances. This is similar to the conclusion obtained in Section 1.6 that showed open-loop controllers to be sensitive to changes in the system model.



UNITED STATES PATENT OFFICE.

ORVILLE WRIGHT AND WILBUR WRIGHT, OF DAYTON, OHIO.

FLYING-MACHINE.

No. 821,393. Specification of Letters Patent. Patented May 22, 1906.

Application filed March 23, 1903. Serial No. 149,220.

lar manner. In either case owing to the varying conditions to be met there are numerous disturbing forces which tend to shift the machine from the position which it should occupy to obtain the desired results. It is the chief object of our invention to provide means for remedying this difficulty, and we will now proceed to describe the construction by means of which these results are accomplished.

Figure 1.11 The Wright brothers' 1903 patent [WW06].

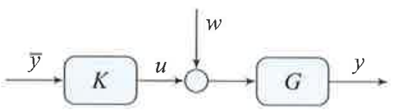


Figure 1.12 Open-loop configuration with input disturbance w.

In closed-loop, Fig. 1.13, we calculate that

$$e = \bar{y} - y = \bar{y} - G(Ke + w) \qquad \Longrightarrow \qquad (1 + GK)e = \bar{y} - Gw$$

and the tracking error is

$$e = \frac{1}{1 + GK}\bar{y} - \frac{G}{1 + GK}w. \tag{1.4}$$

The control gain, K, shows up in both transfer-functions from the inputs, w and $\bar{y}$, to the tracking error, e. High control gains reduce both terms at the same time. That is, the closed-loop solution achieves good tracking *and* rejects the disturbance. This is a most welcome feature and often the main reason for using feedback in control systems. By another coincidence, the coefficient of the first term in (1.4) is the same as the sensitivity function, $S(G)$, calculated in Section 1.6.

Problems

1.1 For each block-diagram in Fig. 1.14 identify *inputs*, *outputs*, and other relevant *signals*, and also describe what physical quantities the signals could represent. Determine whether the system is in *closed-loop* or *open-loop* based on the presence or absence of *feedback*. Is the relationship between the inputs and outputs dynamic or static? Write a simple equation for each block if possible. Which signals are *disturbances*?

1.2 Sketch block-diagrams that can represent the following phenomena as *systems*:

(a) skin protection from sunscreen;
(b) money in a savings account;
(c) a chemical reaction.

Identify potential *input* and *output* signals that could be used to identify cause–effect relationships. Discuss the assumptions and limitations of your *model*. Is the relationship

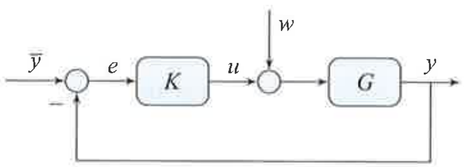


Figure 1.13 Closed-loop feedback configuration with input disturbance w.

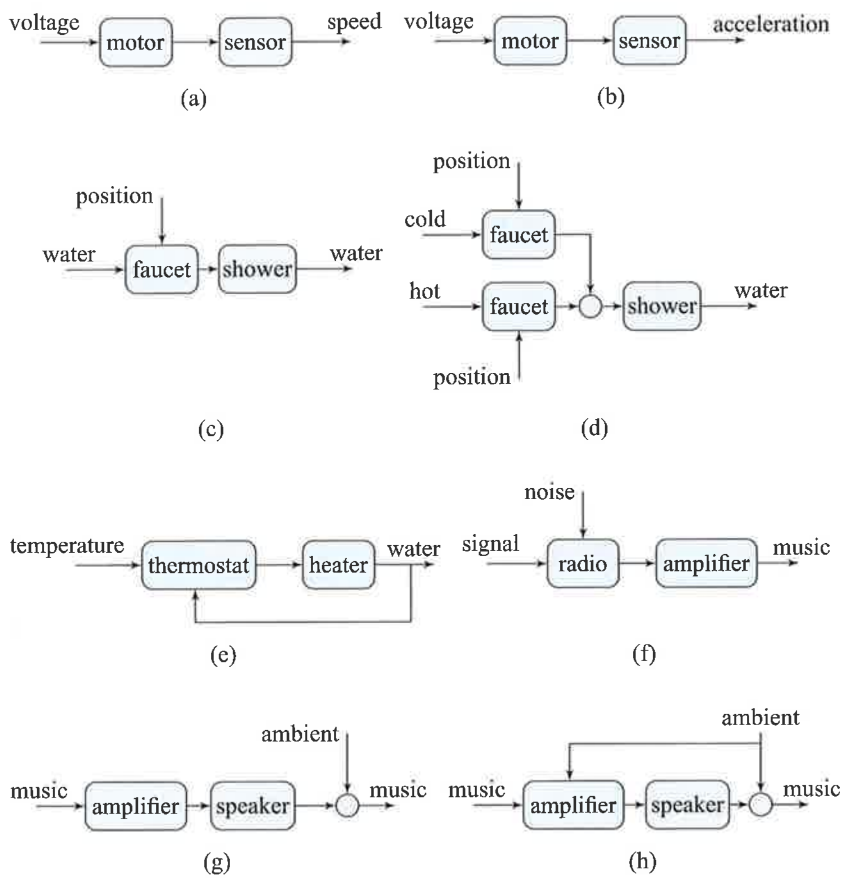


Figure 1.14 Block diagrams for P1.1.

between the inputs and outputs dynamic or static? Write simple equations for each block if possible.

1.3 Mammals are able to regulate their body temperature near 36.5 °C (~98 °F) despite fluctuations in the ambient temperature. Sketch a block-diagram that could represent a possible temperature control system in mammals. Identify disturbances, signals, and possible principles of sensing and actuation that could be used by mammals to lower or increase the body temperature. Compare possible *open-loop* and *closed-loop* solutions. Discuss the difficulties that need to be overcome in each case.

1.4 Most cars are equipped with an *anti-lock braking system* (ABS), which is designed to prevent the wheels from *locking up* when the driver actuates the brake pedal. It helps with emergencies and adverse road conditions by ensuring that traction is maintained on all wheels throughout breaking. An ABS system detects locking of a wheel by comparing the rotational speeds among wheels and modifies the pressure on the hydraulic brake actuator as needed. Sketch a block-diagram that could represent the signals and systems involved in ABS.

1.5 Humans learn to balance standing up early in life. Sketch a block-diagram that represents signals and systems required for standing up. Is there a sensor involved? Actuator? Feedback?

1.6 Sketch a block-diagram that represents the signals and systems required for a human to navigate across an unknown environment. Is there a sensor involved? Actuator? Feedback?

1.7 Repeat P1.5 and P1.6 from the perspective of a blind person.

1.8 Repeat P1.5 and P1.6 from the perspective of a robot or an autonomous vehicle.

1.9 For each block-diagram in Fig. 1.15 compute the transfer-function from the input u to the output y assuming that all blocks are linear.

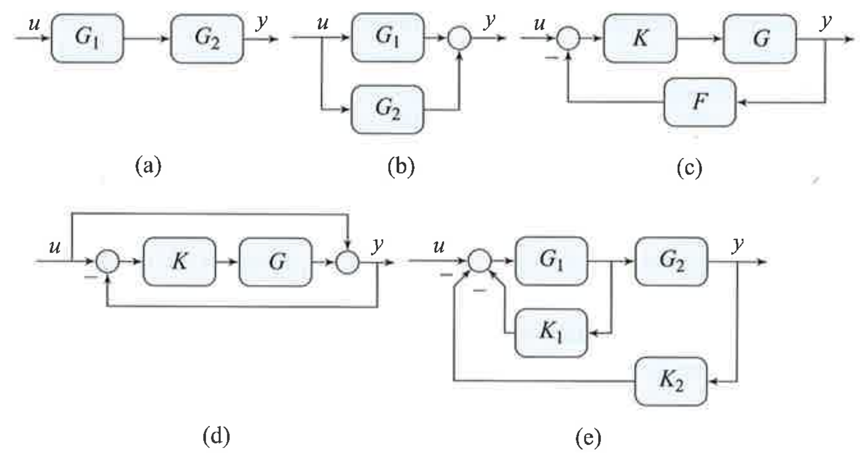


Figure 1.15 Block diagrams for P1.9.

1.10 Students participating in Rice University's Galileo Project [Jen14] set out to carefully reproduce some of Galileo's classic experiments. One was the study of projectile motion using an inclined plane, in which a ball accelerates down a plane inclined at a certain angle then rolls in the horizontal direction with uniform motion for a short while until falling off the edge of a table, as shown in Fig. 1.16. The distance the ball rolled along the inclined plane, ℓ in feet, and the distance from the end of the table to the landing site of the ball, d in inches, were recorded. Some of their data, five trials at two different angles, is reproduced in Table 1.2. Use MATLAB to plot and visualize the data. Fit simple equations, e.g. linear, quadratic, etc., to the data to relate the fall

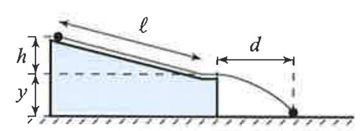


Figure 1.16 Galileo's inclined plane.

height, h, to the horizontal travel distance, d, given in Table 1.2. Justify your choice of equations and comment on the quality of the fit obtained in each case. Estimate using the given data the vertical distance y. Can you also estimate gravity?

Table 1.2 Data for P1.10

	Ramp distance at 13.4°			
Try	1 ft	2 ft	4 ft	6 ft
1	$13\frac{15}{16}$	$19\frac{13}{16}$	$27\frac{11}{16}$	$33\frac{3}{8}$
2	$13\frac{7}{8}$	$19\frac{13}{16}$	$27\frac{3}{4}$	$33\frac{5}{16}$
3	$14\frac{1}{16}$	$19\frac{13}{16}$	$27\frac{3}{4}$	$33\frac{3}{16}$
4	14	$19\frac{3}{4}$	$27\frac{9}{16}$	$33\frac{7}{16}$
5	$13\frac{15}{16}$	$19\frac{3}{4}$	$27\frac{9}{16}$	$33\frac{5}{8}$

	Ramp distance at 6.7°				
Try	1 ft	2 ft	4 ft	6 ft	8 ft
1	$10\frac{11}{6}$	$14\frac{1}{2}$	$20\frac{3}{4}$	$25\frac{7}{16}$	$29\frac{5}{8}$
2	$10\frac{11}{6}$	$14\frac{9}{16}$	$20\frac{3}{4}$	$25\frac{1}{2}$	$29\frac{1}{2}$
3	$10\frac{11}{6}$	$14\frac{1}{2}$	$20\frac{3}{4}$	$25\frac{3}{4}$	$29\frac{1}{2}$
4	$10\frac{11}{6}$	$14\frac{1}{2}$	$20\frac{3}{4}$	$25\frac{1}{2}$	$29\frac{5}{16}$
5	$10\frac{11}{6}$	$14\frac{9}{16}$	$20\frac{3}{16}$	$25\frac{5}{8}$	$29\frac{1}{2}$

1.11 Use MATLAB to determine the parameters α, β, and γ that produce the least-squares fit of the data in Fig. 1.4 to the curves $y(u) = \alpha \tan^{-1}(\beta u)$ and $y(u) = \gamma\, u$. Compare your answers with Figs. 1.5 and 1.6.

1.12 Compute the first-order Taylor series expansion of the function $y(u) = \alpha \tan^{-1}(\beta u)$ about $u = 0$ and use the solution to P1.11 to verify the value of the slope shown in Fig. 1.6.

1.13 Show that the sensitivity function in (1.3) is the one associated with the closed-loop transfer-function $H(G) = GK(1 + GK)^{-1}$.

2 Dynamic Systems

In Chapter 1 we contemplated solutions to our first control problem, a much simplified cruise controller for a car, without taking into account possible effects of *time*. System and controller models were *static* relations between the signals: the output signal, y, the input signal, u, the reference input, $\bar{y}$, and then the disturbance, w. Signals in block-diagrams flow *instantaneously*, and closed-loop solutions derived from such block-diagrams were deemed reasonable if they could be implemented *fast enough*. We drew encouraging conclusions from simple analysis but no rationale was given to support the conclusions if time were to be taken into consideration.

Of course, it is perfectly fine to construct a static mathematical model relating a car's pedal excursion with its terminal velocity, as long as we understand the model setup. Clearly a car does not reach its terminal velocity *instantaneously*! If we expect to implement the feedback cruise controller in a real car, we have to be prepared to say what happens between the time at which a terminal target velocity is set and the time at which the car reaches its terminal velocity. Controllers have to *understand* that it takes *time* for the car to reach its terminal velocity. That is, we will have to incorporate time not only into models and tools but also into controllers. For this reason we need to learn how to work with *dynamic* systems.

In the present book, mathematical models for dynamic systems take the form of ordinary[1] differential equations where signals evolve continuously in time. Bear in mind that this is not a course on differential equations, and previous exposure to the mathematical theory of differential equations helps. Familiarity with material covered in standard text books, e.g. [BD12], is enough. We make extensive use of the Laplace transform and provide a somewhat self-contained review of relevant facts in Chapter 3.

These days, when virtually all control systems are implemented in some form of digital computer, one is compelled to justify why not to discuss control systems directly from the point of view of discrete-time signals and systems. One reason is that continuous-time signals and systems have a long tradition in mathematics and physics that has established a *language* that most scientists and engineers are accustomed to. The converse is unfortunately not true, and it takes time to get comfortable with interpreting discrete-time models and making sense of some of the implied assumptions that come with them, mainly the effects of sampling and related practical issues, such as

[1] Ordinary, as opposed to *partial*, means that derivatives appear only with respect to one variable; in our case, time.

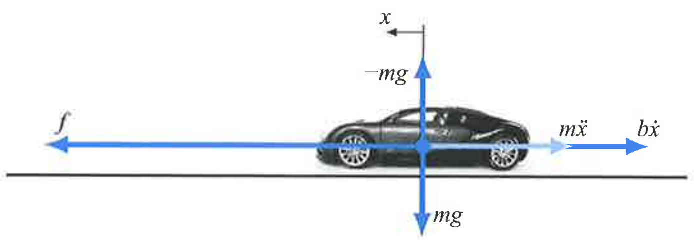


Figure 2.1 Free-body diagram showing forces acting in a car.

quantization and aliasing. In fact, for physical systems, it is impossible to appropriately choose an adequate sampling rate without having a good idea of the continuous-time model of the system being controlled. Finally, if a system is well modeled and a controller is properly designed in continuous-time, implementation in the form of a discrete-time controller is most of the time routine, especially when the available hardware sampling rates are fast enough.

2.1 Dynamic Models

Let us start with some notation: we denote time by the real variable t in the interval $[0, \infty)$, where 0 can be thought of as an arbitrary origin of time before which we are not interested in the behavior of the system or its signals. We employ functions of real variables to describe signals and use standard functional notation to indicate the dependence of signals on time. For example, the dynamic signals y and u are denoted as $y(t)$ and $u(t)$. At times, when no confusion is possible, we omit the dependence of signals on t.

We claimed in Chapter 1 that models should be rooted in well-defined experiments. Planning and performing experiments for dynamic systems is a much more complex task, which we do not have room to address in detail here. Instead, we will reach out to physics to help us introduce an abstract dynamic model, the parameters of which will later be determined through experiments.

In the tradition of simplified physical modeling, we use Newton's law to write equations for a car based on the free-body diagram shown in Fig. 2.1. The car is modeled as a particle with mass $m > 0$ and, after balancing all forces in the x-direction, we obtain the differential equation

$$m\ddot{x}(t) + b\dot{x}(t) = f(t),$$

where x is the linear coordinate representing the position of the car, $b \geq 0$ is the coefficient of friction, and f is a force, which we will use to put the car into motion. Much can be argued about the exact form of the friction force, which we have assumed to be *viscous*, that is of the form $-b\,v(t)$, linear, and opposed to the velocity $v(t) = \dot{x}(t)$. As we are interested in modeling the velocity of the car and not its position, it is convenient to rewrite the differential equation in terms of the velocity v, obtaining

$$m\dot{v}(t) + bv(t) = f(t). \qquad (2.1)$$

Figure 2.2 Block-diagram with integrator.

In order to complete our model we need to relate the car driving force, f, to the pedal excursion, u. Here we could resort to experimentation or appeal to basic principles. With simplicity in mind we choose a linear static model:

$$f = pu, \qquad (2.2)$$

where p represents a *pedal gain*, which can be determined experimentally by methods similar to the ones used in Chapter 1.

Of course, no one should believe that the simple force model (2.2) can accurately represent the response of the entire powertrain of the car in a variety of conditions. Among other things, the powertrain will have its own complex dynamic behaviors, which (2.2) gracefully ignores. Luckily, the validity[2] of such simplification depends not only on the behavior of the actual powertrain but also on the purpose of the model. In many cases, the time-constants[3] of the powertrain are much *faster* than the time-constant due to the inertial effects of the entire car. In this context, a simplified model can lead to satisfactory or at least insightful results when the purpose of the model is, say, predicting the velocity of the car. A human driver certainly does not need to have a deep knowledge of the mechanical behavior of an automobile for driving one!

Combining Equations (2.1) and (2.2), and labeling the velocity as the output of the system, i.e. $y(t) = v(t)$, we obtain the differential equation

$$\dot{y}(t) + \frac{b}{m}y(t) = \frac{p}{m}u(t), \qquad (2.3)$$

which is the mathematical dynamic model we will use to represent the car in the dynamic analysis of the cruise control problem.

2.2 Block-Diagrams for Differential Equations

In Chapter 1 we used block-diagrams to represent the interaction of signals, systems, and controllers. If we interpret differential equations as a relationship between signals and the signals' derivatives, it should be no surprise that ordinary differential equations can be represented in block-diagrams. The key is to use an *integrator* to relate the derivative to its primitive. An integrator block should be able to produce at its output the integral of its input. Alternatively, we can see the integrator input as the derivative of its output signal, as depicted in Fig. 2.2. We will study in Chapter 5 a number of physical devices that can be used to physically implement integrators.

Assuming that integrator blocks are available, all that is left to do is to rewrite the ordinary differential equation, isolating its highest derivative. For example, we

[2] A better word here may be *usefulness*. [3] More about that soon!

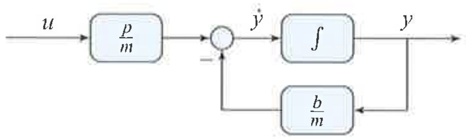


Figure 2.3 Dynamic model of the car: m is the mass, b is the viscous friction coefficient, p is the pedal gain, u is the pedal excursion, and y is the car's velocity.

rewrite (2.3) as

$$\dot{y}(t) = \frac{p}{m}u(t) - \frac{b}{m}y(t),$$

which can be represented by the block-diagram in Fig. 2.3.

Note the presence of a *feedback loop* in the diagram of Fig. 2.3! For this reason, tools for analyzing feedback loops often draw on the theory of differential equations and vice versa. We will explore the realization of differential equations using block-diagrams with integrators in detail in Chapter 5.

2.3 Dynamic Response

The differential equation (2.3) looks very different from the static linear models considered earlier in Chapter 1. In order to understand their differences and similarities we need to understand how the model (2.3) *responds* to inputs. Our experiment in Section 1.1 consisted of having a constant pedal excursion and letting the car reach a terminal velocity. We shall first attempt to emulate this setup using the differential equation (2.3) as a model.

A constant pedal excursion corresponds to the input function

$$u(t) = \tilde{u}, \quad t \geq 0,$$

where $\tilde{u}$ is constant. In response to this input we expect that the solution of the differential equation (2.3) approaches a constant terminal velocity, $\tilde{y}$, after some time has passed. The value of this terminal velocity can be calculated after noticing that

$$y(t) = \tilde{y}, \quad t \geq T \qquad \Longrightarrow \qquad \dot{y}(t) = 0, \quad t > T,$$

in which case (2.3) reduces to

$$\frac{b}{m}\tilde{y} = \frac{p}{m}\tilde{u} \qquad \Longrightarrow \qquad \tilde{y} = \frac{p}{b}\tilde{u}.$$

It is this relation that should be compared with the static model developed earlier. Experiments similar to the ones in Section 1.1 can be used to determine the value of the ratio p/b. In the language of differential equations the function

$$y_P(t) = \tilde{y} = \frac{p}{b}\tilde{u}$$

is a *particular solution* to the differential equation (2.3). See [BD12] for details.

The particular solution cannot, however, be a complete solution: if the initial velocity of the car at time $t = 0$ is not equal to $\tilde{y}$ then $y(0) = y_0 \neq \tilde{y} = y_P(0)$. The remaining component of the solution is found by solving the *homogeneous* version of Equation (2.3):

$$\dot{y}(t) + \frac{b}{m}y(t) = 0.$$

All solutions to this equation can be shown to be of the form

$$y_H(t) = e^{\lambda t}, \tag{2.4}$$

where the constant λ is determined upon substitution of $y_H(t)$ into (2.3):

$$\dot{y}_H(t) + \frac{b}{m}y_H(t) = \left(\lambda + \frac{b}{m}\right)e^{\lambda t} = 0.$$

This is an algebraic equation that needs to hold for all $t \geq 0$, in particular $t = 0$, which will happen only if λ is a zero of the *characteristic equation*:

$$\lambda + \frac{b}{m} = 0 \qquad \Longrightarrow \qquad \lambda = -\frac{b}{m}. \tag{2.5}$$

The complete solution to Equation (2.3) is a combination of the particular solution, $y_P(t)$, with all possible solutions to the homogeneous equation:[4]

$$y(t) = y_P(t) + \beta\, y_H(t) = \tilde{y} + \beta\, e^{\lambda t},$$

in which the constant β is calculated so that $y(t)$ matches the *initial condition*, y_0, at $t = 0$. That is,

$$y(0) = \tilde{y} + \beta = y_0 \qquad \Longrightarrow \qquad \beta = y_0 - \tilde{y}.$$

Putting it all together, the complete response is

$$y(t) = \tilde{y}\left(1 - e^{\lambda t}\right) + y_0\, e^{\lambda t}, \qquad t \geq 0, \tag{2.6}$$

where

$$\lambda = -\frac{b}{m}, \qquad \tilde{y} = \frac{p}{b}\tilde{u}. \tag{2.7}$$

Plots of $y(t)$ for various values of λ and y_0 are shown in Fig. 2.4 for $\tilde{y} = 1$. Note how the responses *converge* to $\tilde{y}$ for all negative values of λ. The more negative the value of λ, the faster the convergence. When λ is positive the response does not converge to $\tilde{y}$, even when y_0 is very close to $\tilde{y}$.

It is customary to evaluate how fast the solution of the differential equation (2.3) converges to $\tilde{y}$ by analyzing its response to a zero initial condition $y_0 = 0$ and a nonzero $\tilde{y} \neq 0$. This is known as a *step response*. When $\lambda < 0$, the constant

$$\tau = -\frac{1}{\lambda} \tag{2.8}$$

[4] In this simple example there is only one such solution, $y_H(t)$ given in (2.4).

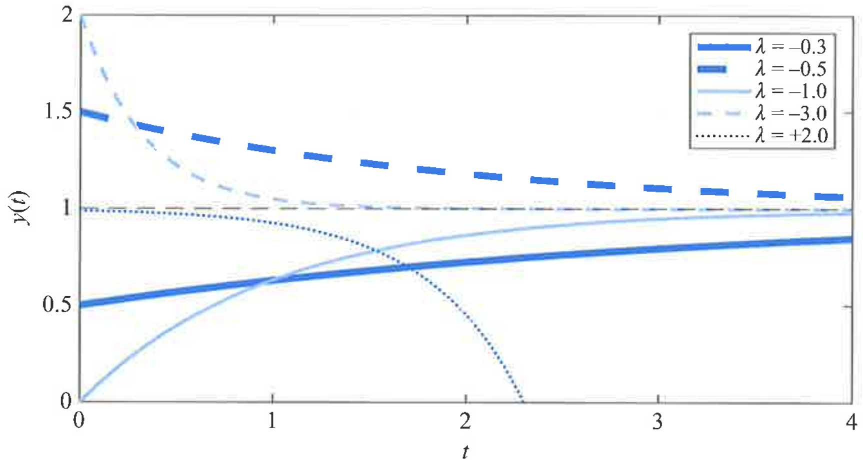


Figure 2.4 Plots of $y(t) = \bar{y}\left(1 - e^{\lambda t}\right) + y_0 e^{\lambda t}$, $t \geq 0$, with $\bar{y} = 1$; λ and y_0 are as shown.

is the *time-constant*. In P2.1 you will show that τ has units of time. At select times $t = \tau$ and $t = 3\tau$,

$$y(\tau) = \bar{y}\left(1 - e^{-1}\right) \approx 0.63\bar{y}, \qquad y(3\tau) = \bar{y}\left(1 - e^{-3}\right) \approx 0.95\bar{y}.$$

As τ depends only on λ and not on $\bar{y}$, it is possible to compare the rate of convergence of different systems modeled by linear ordinary differential equations by comparing their time-constants. For differential equations more complex than (2.3), the time-constant is *defined* as the time it takes the step response to reach 63% of its terminal value. The smaller the time-constant, the faster the convergence.

Another measure of the rate of change of the response is the *rise-time*, t_r, which is the time it takes the step response to go from 10% to 90% of its final value.[5] Calculating

$$y(t_1) = \bar{y}\left(1 - e^{\lambda t_1}\right) = 0.1\bar{y}, \qquad y(t_2) = \bar{y}\left(1 - e^{\lambda t_2}\right) = 0.9\bar{y},$$

we obtain

$$t_r = t_2 - t_1 = \ln(1/9)\lambda^{-1} = \ln(9)\tau \approx 2.2\,\tau. \tag{2.9}$$

Again, the smaller the rise-time, the faster the convergence.

2.4 Experimental Dynamic Response

As shown in Section 2.3, the terminal velocity attained by the dynamic linear model, the differential equation (2.3), is related to the static linear model, the algebraic equation (1.1), through $\gamma = \bar{y}/\bar{u} = p/b$. This means that the ratio p/b can be determined in the same way as was done in Section 1.1. A new experiment is needed to determine the parameter $\lambda = -b/m$, which does not influence the terminal velocity but affects the rate at which the car approaches the terminal velocity.

[5] For systems in which the output develops only after a delay it is easier to measure the rise-time than the time-constant.

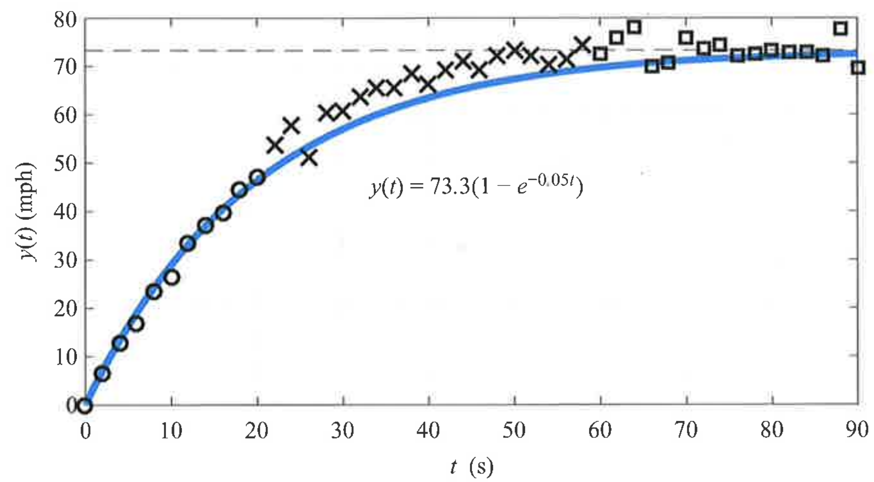


Figure 2.5 Experimental velocity response of a car to a constant pedal excursion, $u(t) = \bar{u} = 1$ in, $t \geq 0$; samples are marked as circles, crosses, and squares.

First let us select a velocity around which we would like to *build* our model, preferably a velocity close to the expected operation of the cruise controller. Looking at Figs. 1.4 through 1.6, we observe that a pedal excursion of around 1 in will lead to a terminal velocity around 70 mph, which is close to highway speeds at which a cruise controller may be expected to operate. We perform the following dynamic experiment: starting at rest, apply constant pedal excursion, $u(t) = \bar{u} = 1$ in, $t \geq 0$, and collect samples of the instantaneous velocity until the velocity becomes approximately constant. In other words, perform an *experimental step response*. The result of one such experiment may look like the plot in Fig. 2.5, in which samples (marked as circles, crosses, and squares) have been collected approximately every 2 s for 90 s.

We proceed by *fitting* the data in Fig. 2.5 to a function like (2.6) where the initial condition, y_0, is set to zero before estimating the parameters $\bar{y}$ and λ. This fit can be performed in many ways. We do it as follows: we first average the samples over the last 30 s in Fig. 2.5 (squares) to compute an estimate of the terminal velocity. From the data shown in Fig. 2.5 we obtain the estimate $\bar{y} \approx 73.3$ mph. If $y(t)$ is of the form (2.6) then

$$r(t) = 1 - y(t)/\bar{y} = e^{\lambda t} \qquad \Longrightarrow \qquad \ln r(t) = \lambda t.$$

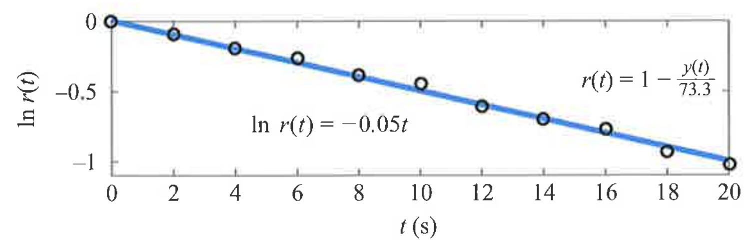


Figure 2.6 Plot of $\ln r(t)$ and fitted linear model.

That is, $\ln r(t)$ is a line with slope λ. With this in mind we plot $\ln r(t)$ in Fig. 2.6 using samples taken from the first 20 s from Fig. 2.5 (circles) and estimate the slope of the line $\ln r(t)$, that is $\lambda = -0.05$. The parameters b/m and p/m are then estimated based on the relationships

$$\begin{aligned} \frac{b}{m} &= -\lambda \approx 0.05\,\text{s}^{-1}, \\ \frac{p}{b} &= \frac{\bar{y}}{\bar{u}} \approx 73.3\,\text{mph/in}, \\ \frac{p}{m} &= \frac{b}{m} \times \frac{p}{b} = -\lambda \times \frac{\bar{y}}{\bar{u}} = 3.7\,\text{mph/(in s)}. \end{aligned} \tag{2.10}$$

Note that this model has a static gain of about 73.3 mph/in which lies somewhere between the two static linear models estimated earlier in Fig. 1.6. Indeed, this is the intermediate gain value that was used in Section 1.5 to calculate one of the static closed-loop transfer-functions in Table 1.1.

The estimation of the structure and the parameters of a dynamic system from experiments is known as *system identification*. The interested reader is referred to [Lju99] for an excellent introduction to a variety of useful methods.

2.5 Dynamic Feedback Control

We are now ready to revisit the feedback solution proposed in Section 1.3 for solving the cruise control problem. Let us keep the structure of the feedback loop the same, that is let the *proportional controller*

$$u(t) = Ke(t), \qquad e(t) = \bar{y} - y(t) \tag{2.11}$$

be connected as in Fig. 1.8. Note a fundamental difference between this controller and the one analyzed before: in Section 1.5 the signals e and y were the terminal error and terminal velocity; controller (2.11) uses the *dynamic* error signal $e(t)$ and velocity $y(t)$. This dynamic feedback loop can be practically implemented if a sensor for the instantaneous velocity, $y(t)$, is used. Every vehicle comes equipped with one such sensor, the speedometer.[6]

In order to analyze the resulting dynamic feedback control loop we replace the system model, G, with the dynamic model, the differential equation (2.3), to account for the car's dynamic response to changes in the pedal excursion. In terms of block-diagrams, we replace G in Fig. 1.8 by the block-diagram representation of the differential equation (2.3) from Fig. 2.3. The result is the block-diagram shown in Fig. 2.7. Using Equations (2.3) and (2.11) we eliminate the input signal, $u(t)$, to obtain

$$\dot{y}(t) + \left(\frac{b}{m} + \frac{p}{m}K\right) y(t) = \frac{p}{m}K\bar{y}. \tag{2.12}$$

[6] The speedometer measures the speed but it is easy to infer the direction, hence the velocity, in this simple one-dimensional setup.

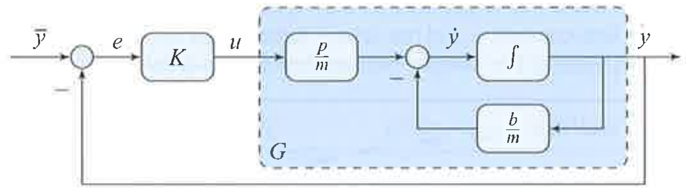


Figure 2.7 Dynamic closed-loop connection of the car model with proportional controller.

This linear ordinary differential equation governs the behavior of the closed-loop system. In the next chapters, you will learn to interpret this equation in terms of a closed-loop transfer-function using the Laplace transform. For now we proceed in the time domain and continue to work with differential equations.

Since Equation (2.12) has the same structure as Equation (2.3), its solution is also given by (2.6). That is

$$y(t) = \tilde{y}\left(1 - e^{\lambda t}\right) + y_0 e^{\lambda t}, \qquad t \geq 0,$$

with the constants

$$\lambda = -\frac{b}{m} - \frac{p}{m}K, \qquad \tilde{y} = \frac{(p/m)K}{(b/m) + (p/m)K}\bar{y} = \frac{(p/b)K}{1 + (p/b)K}\bar{y}.$$

When K is positive, the decay-rate λ is negative and hence $y(t)$ converges to $\tilde{y}$. In terms of the target velocity $\bar{y}$ we can write

$$\lim_{t\to\infty} y(t) = \tilde{y} = H(0)\bar{y}, \qquad H(0) = \frac{(p/b)K}{1 + (p/b)K}. \tag{2.13}$$

We refer to components of the response of a dynamic system that persist as the time grows large as steady-state solutions. In this case, the closed-loop has a constant steady-state solution

$$y_{ss}(t) = \tilde{y} = H(0)\bar{y}.$$

Note that the value of $H(0)$ is equal to the static closed-loop transfer-function computed in Section 1.5 if G is replaced with p/b. As we will see in Chapter 4, this is not a mere coincidence. In terms of *tracking error*,

$$\lim_{t\to\infty} e(t) = \lim_{t\to\infty} \bar{y} - y(t) = S(0)\bar{y}, \qquad S(0) = \frac{1}{1 + (p/b)K}. \tag{2.14}$$

The function $S(0) = 1 - H(0)$ is the static closed-loop *sensitivity function* computed before in Section 1.6. The reason for using the notation $H(0)$ and $S(0)$ will become clear in the next chapters.

For various values of $G = p/b$, including $p/b \approx 73.3$ which we estimated in Section 2.4, the steady-state closed-loop solution will track the reference $\bar{y}$ with accuracy $S(0) = 1 - H(0)$, which can be computed from the values listed in Table 1.1. In steady-state, the closed-loop behaves as predicted by the static analysis in Section 1.3. The dynamic analysis goes a step further: it predicts the rate at which convergence occurs,

Table 2.1 Open- and closed-loop steady-state transfer-function, $H(0)$, steady-state sensitivity, $S(0)$, steady-state limit, $\bar{y}$, time-constant, τ, and rise-time, t_r, calculated for $b/m = 0.05$ and $p/b = 73.3$ and a constant target output of $\bar{y} = 60$ mph. The open-loop solution is from Section 1.4.

K	$H(0)$	$S(0)$	$\tilde{y}$ (mph)	τ (s)	t_r (s)
Open-loop	1.00	0.00	60	20.0	43.9
0.02	0.60	0.40	36	8.1	17.8
0.05	0.79	0.21	47	4.3	9.4
0.50	0.97	0.03	58	0.5	1.2

which is related to the parameters λ and τ:

$$\lambda = -\frac{b}{m} - \frac{p}{m}K, \qquad \tau = -\lambda^{-1} = \frac{m}{b + pK}.$$

Notice that the time-constant, τ, becomes smaller as K grows. A numerical comparison for various values of gain, K, including the open-loop solution[7] (from Section 1.4), is given in Table 2.1. The corresponding dynamic responses calculated from zero initial conditions, $y(0) = 0$, are plotted in Fig. 2.8.

Some numbers in Table 2.1 and Fig. 2.8 look suspicious. Is it really possible to lower time-constants so much? Take, for example, the case of the largest gain $K = 0.5$: here we have almost perfect tracking (3% error) with a closed-loop rise-time that is more than 40 times faster than in open-loop. This kind of performance improvement is unlikely to be achieved by any controller that simply steps into the accelerator pedal. Surely there must be a catch! Indeed, so far we have been looking at the system output, the car's velocity, $y(t)$, and have paid little attention to the control input, the pedal excursion, $u(t)$. We shall now look at the control input in search of clues that could explain the impressive performance of the closed-loop controller.

The control inputs, $u(t)$, associated with the dynamic responses in Fig. 2.8 are plotted in Fig. 2.9. In Fig. 2.9 we see that the feedback controller is injecting into the system, the car, large inputs, pedal excursions, in order to achieve better tracking and faster response. The larger the control gain, K, the larger the required pedal excursion. Note that in this case the maximum required control signal happens at $t = 0$, when the tracking error is at a maximum, and

$$u(0) = Ke(0) = K(\bar{y} - y(0)) = K\bar{y}.$$

Clearly, the larger the gain, K, the larger the control input, u. For instance, with $K = 0.5$ the controller produces an input that exceeds the maximum possible pedal excursion of 3 in, which corresponds to *full throttle*. In other words, the control input is *saturated*. With $K = 0.05$ we have $u(0) = 3$, which is full throttle. Of course, any conclusions drawn for $K > 0.05$ will no longer be valid or, at least, not very accurate. It is not

[7] What we mean by open-loop solution is the simulation of the diagram in Fig. 1.7 where G is replaced by the differential equation model (2.3) but $K = (p/b)^{-1}$ is still the constant open-loop gain as calculated in Section 1.4. We are not yet equipped to speak of G^{-1} as the inverse of a dynamic model, which we will do in Sections 4.7, 5.1, and 8.6.

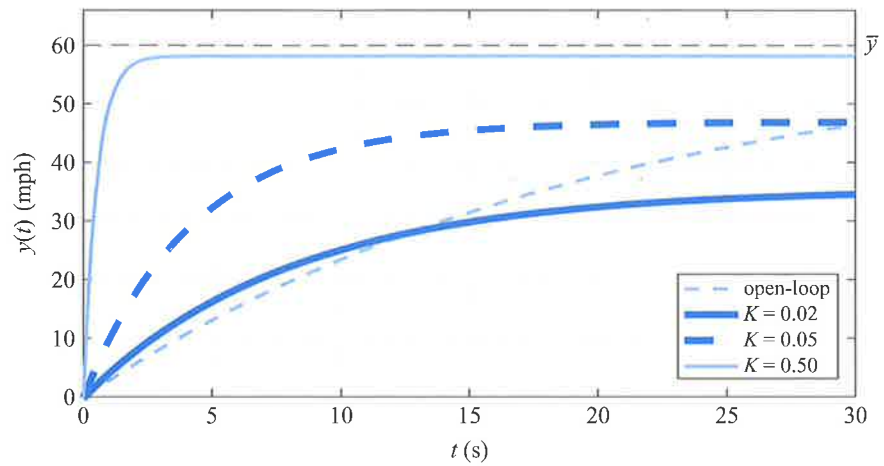


Figure 2.8 Open- and closed-loop dynamic response, $y(t)$, for the linear car velocity model (2.12) calculated for $b/m = 0.05$ and $p/b = 73.3$ and a constant target output of $\bar{y} = 60$ mph with proportional control (2.11) for various values of gain, K; the open-loop solution is from Section 1.4.

possible to achieve some of the predicted ultra-fast response times due to limitations in the system, in this case pedal and engine saturation, that were not represented in the linear models used to design and analyze the closed-loop. Ironically, the gain $K = 0.02$ is one for which the pedal excursion remains well below saturation, and is perhaps the one case in which the (poor) performance predicted by the linear model is likely to be accurate.

2.6 Nonlinear Models

In Section 2.5 we saw controllers that produced inputs that led to saturation of the system input, the car's pedal excursion. In some cases the required control input exceeded full

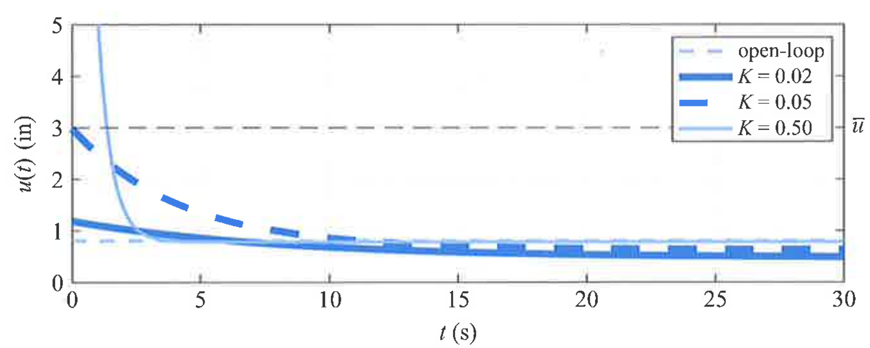


Figure 2.9 Open- and closed-loop control inputs (pedal excursion) corresponding to the dynamic responses in Fig. 2.8; the largest possible pedal excursion is 3 in.

throttle. In this section we digress a little to introduce a simple nonlinear model that can better predict the behavior of the system in closed-loop when saturation is present. This is not a course in nonlinear control, and the discussion will be kept at a very basic level. The goal is to be able to tell what happens in our simple example when the system reaches saturation.

In order to model the effect of saturation we will work with a nonlinear differential equation of the form

$$\dot{y}(t) + \frac{c}{m}\tan(\alpha^{-1}y(t)) = \frac{d}{m}u(t), \qquad u(t) \in [0, 3]. \tag{2.15}$$

When $u(t) = \tilde{u}$ is constant, one particular solution to (2.15) is $y(t) = \tilde{y}$, where

$$\frac{c}{m}\tan(\alpha^{-1}\tilde{y}) = \frac{d}{m}\tilde{u} \qquad \Longrightarrow \qquad \tilde{y} = \alpha\tan^{-1}(\beta\tilde{u}), \quad \beta = \frac{d}{c}.$$

This means that the steady-state response of the nonlinear differential equation (2.15) matches the empirical nonlinear fit performed earlier in Fig. 1.5. Moreover, at least for small values of $y(t)$, we should expect[8] that (see P2.2)

$$c\tan(\alpha^{-1}y(t)) \approx by(t), \qquad b = \alpha^{-1}c.$$

Intuitively, as long as $y(t)$ remains small and $u(t) \in [0, 3]$, the dynamic response of the nonlinear differential equation (2.15) should stay *close* to the dynamic response of the linear differential equation (2.3).

In order to estimate suitable parameters c/m, d/m, and α, we proceed as follows: first we borrow α and β from our previously computed nonlinear static fit ($\alpha = 82.8$ mph/in, $\beta = 1.2\,\text{in}^{-1}$, see Fig. 1.5), then we estimate $b/m = -\lambda$ from the linear dynamic experiment described in Section 2.4, and calculate

$$\begin{aligned}\frac{c}{m} &= \alpha \times \frac{b}{m} = 82.8 \times 0.05 = 4.1\,\text{mph/in},\\ \frac{d}{m} &= \beta \times \frac{c}{m} = 1.2 \times 4.1 = 5.0\,\text{mph/in s}.\end{aligned} \tag{2.16}$$

The resulting nonlinear model has a steady-state solution that matches the static fit from Fig. 1.5 and a time-constant close to that of the linear model from Section 2.1.

It is in closed-loop, however, that the nonlinear model will likely expose serious limitations of the controller based on the linear model (2.3). In order to capture the limits on pedal excursion we introduce the saturation nonlinearity:

$$\text{sat}_{(\underline{u},\overline{u})}(u(t)) = \begin{cases}\overline{u}, & u(t) > \overline{u},\\ u(t), & \underline{u} \le u(t) \le \overline{u},\\ \underline{u}, & u(t) < \underline{u}.\end{cases}$$

[8] This notion will be formalized in Section 5.4.

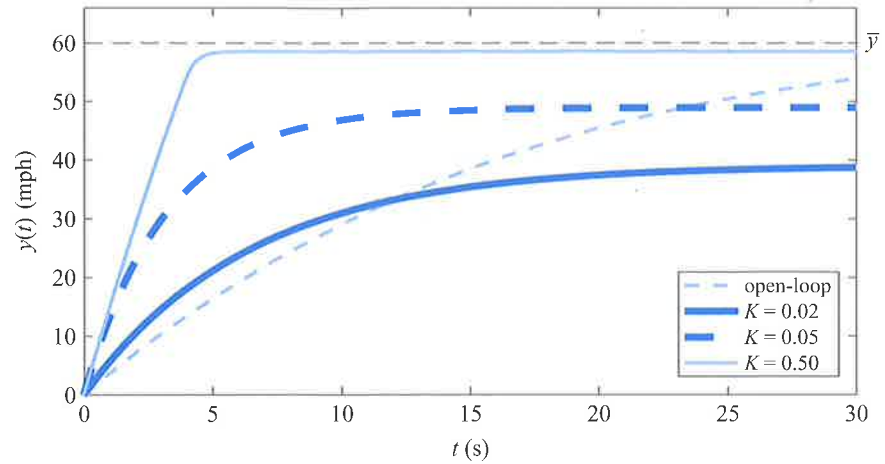


Figure 2.10 Open- and closed-loop dynamic response, $y(t)$, produced by the nonlinear car velocity model (2.17) calculated with $\alpha = 1.2$, $\beta = 82.8$, $c/m = 4.1$, and $d/m = 5.0$ and a constant target output of $\bar{y} = 60$ mph under proportional control (2.11) and various values of gain, K. Compare this with Fig. 2.8.

In the car model,[9] $\underline{u} = 0$ and $\overline{u} = 3$ in. The complete nonlinear model is

$$\dot{y}(t) + \frac{c}{m}\tan(\alpha^{-1}y(t)) = \frac{d}{m}\text{sat}_{(0,3)}(u(t)), \qquad u(t) = K(\bar{y} - y(t)).$$

Eliminating $u(t)$ we obtain

$$\dot{y}(t) + \frac{c}{m}\tan(\alpha^{-1}y(t)) = \frac{d}{m}\text{sat}_{(0,3)}(K(\bar{y} - y(t))). \tag{2.17}$$

The above nonlinear ordinary differential equations cannot be solved analytically but can be simulated using standard numerical integration methods, e.g. one of the Runge–Kutta methods [BD12].

In order to see the effect of the nonlinearities on the closed-loop performance we repeat the simulations performed in Figs. 2.8 and 2.9, this time using the nonlinear feedback model (2.17). We show in Figs. 2.10 and 2.11 the nonlinear system closed-loop response, $y(t)$, and the control input, $u(t)$, for various values of the gain, K. These should be compared with Figs. 2.8 and 2.9. For values of $K = 0.05$ and $K = 0.02$, when the control input predicted using the linear model is within the linear region, i.e. $u(t) \in [0, 3]$, and the speeds are small, the nonlinearity has a minor impact. However, in the case of the larger gain $K = 0.5$, where the control input is heavily saturated (see Fig. 2.11), the response is significantly different. In particular, the extraordinarily fast response predicted by the linear model is not realized in the nonlinear model. In this simple example, the slower response seems to be the only apparent consequence. This will not always be the case, and severe nonlinearities will often negatively affect the performance of closed-loop systems. See Section 5.8.

[9] Note that by setting $\underline{u} = 0$ we prevent the model from applying any braking force.

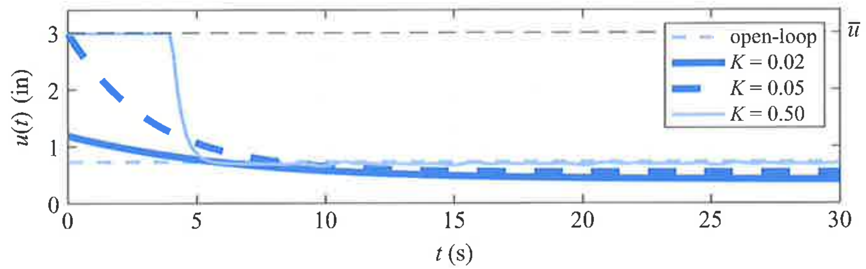


Figure 2.11 Open- and closed-loop control input, pedal excursion, $u(t)$, produced by the car velocity nonlinear model equation (2.17) under proportional control (2.11); the largest possible pedal excursion is 3 in; note the marked effect of pedal saturation in the case of the highest gain $K = 0.5$ and its impact in Fig. 2.10.

2.7 Disturbance Rejection

We now return to linear models to talk about a much desired feature of feedback control: *disturbance rejection*. Consider a modified version of the cruise control problem where the car is on a slope, as illustrated in Fig. 2.12. Newton's law applied to the car produces the differential equation[10]

$$m\dot{y}(t) + by(t) = f(t) - mg\sin(\theta(t)),$$

where $y = \dot{x}$ is the velocity of the car, θ is the angle the slope makes with the horizontal, and g is the gravitational acceleration.[11] When $\theta = 0$ the car is on the flat, and the model reduces to (2.1). As before, adoption of a linear model for the relationship between the force, f, and the pedal excursion, u, i.e. $f = pu$ from (2.2), produces the differential equation

$$\dot{y}(t) + \frac{b}{m}y(t) = \frac{p}{m}u(t) - g\sin(\theta(t)).$$

This is a linear differential equation except for the way in which $\theta(t)$ enters the equation. In most cases, the *signal* $\theta(t)$ is not known ahead of time, and can be seen as a *nuisance* or *disturbance*. Instead of working directly with θ, it is convenient to introduce the disturbance signal

$$w(t) = -\frac{mg}{p}\sin(\theta(t)) \qquad (2.18)$$

[10] Strictly speaking, this differential equation is true only if $\theta(t)$ is constant. When $\theta(t)$ is not constant, the car is subject to a (non-working) force that originates from changes in its frame of reference. This additional force can itself be treated as an additional disturbance if the changes in slope are moderate. No rollercoasters please!

[11] As we insist on using a non-standard unit for measuring velocity (mph), g will have to be expressed in mph/s, or $g = 9.8\,\text{m/s}^2 \approx 9.8 \times 3600/1609 \approx 21.9$ mph/s. Ugh! That is ugly!

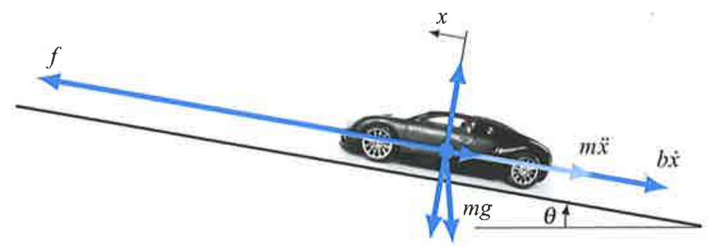


Figure 2.12 Free-body diagram showing forces acting on a car on a road slope.

as affecting the linear model

$$\dot{y}(t) + \frac{b}{m}y(t) = \frac{p}{m}u(t) + \frac{p}{m}w(t). \qquad (2.19)$$

This differential equation is linear and can be analyzed with simpler tools.

The car model with the input disturbance is represented in closed-loop by the block-diagram in Fig. 2.13, which corresponds to the equations

$$\dot{y}(t) + \frac{b}{m}y(t) = \frac{p}{m}u(t) + \frac{p}{m}w(t), \qquad u(t) = K(\bar{y} - y(t)),$$

or, after eliminating u,

$$\dot{y}(t) + \left(\frac{b}{m} + \frac{p}{m}K\right)y(t) = \frac{p}{m}K\bar{y} + \frac{p}{m}w(t). \qquad (2.20)$$

In order to understand how disturbances affect the closed-loop behavior we shall analyze the following scenario: suppose that the car is traveling on the flat, $w = \theta = 0$, with the cruise control in closed-loop at the steady-state velocity $y_{ss}(t) = H(0)\bar{y}$. At time $t = 0$ s, the car hits a 10% grade slope, $\bar{\theta} \approx 5.7°$. We use (2.18) to calculate the disturbance $w(t) = \bar{w} \approx -0.26$, $t \geq 0$.

The dynamic response of the car to the change in slope can be computed by formula (2.6) after setting

$$y_0 = H(0)\bar{y}, \qquad w(t) = \bar{w}, \qquad t \geq 0$$

and calculating

$$\lambda = -\frac{b}{m} - \frac{p}{m}K, \qquad \tilde{y} = \frac{(p/m)K\bar{y} + (p/m)\bar{w}}{b/m + (p/m)K}.$$

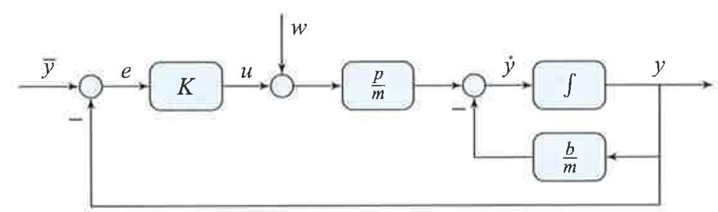


Figure 2.13 Closed-loop connection of the car showing the slope disturbance $w = -(mg/p)\sin(\theta)$.

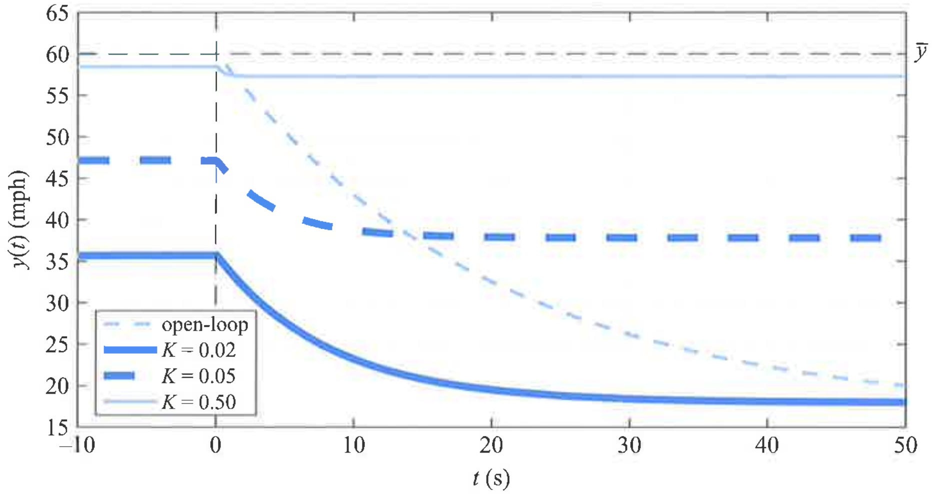


Figure 2.14 Closed-loop response of the velocity of the car with proportional cruise control (linear model (2.12), $b/m = 0.05$ and $p/b = 73.3$) to a change in road slope at $t = 0$, from flat to 10% grade for various values of the control gain.

It is useful to split $\tilde{y}$ into two components:

$$\tilde{y} = H(0)\bar{y} + D(0)\bar{w}, \qquad D(0) = \frac{p}{b}\frac{1}{1 + (p/b)K}, \tag{2.21}$$

where $H(0)$ is the same as in (2.13), and with which (2.6) becomes

$$y(t) = \tilde{y}\left(1 - e^{\lambda t}\right) + y_0 e^{\lambda t} = H(0)\bar{y} + \left(1 - e^{\lambda t}\right) D(0)\bar{w}.$$

The closed-loop response and the open-loop response are plotted in Fig. 2.14 for various values of the gain, K. The predicted change in velocity is equal to

$$\Delta y(t) = y(t) - y_0 = y(t) - H(0)\bar{y} = \left(1 - e^{\lambda t}\right) D(0)\bar{w}.$$

From (2.21), the larger K, the smaller $D(0)$, hence the smaller the change in velocity induced by the disturbance.

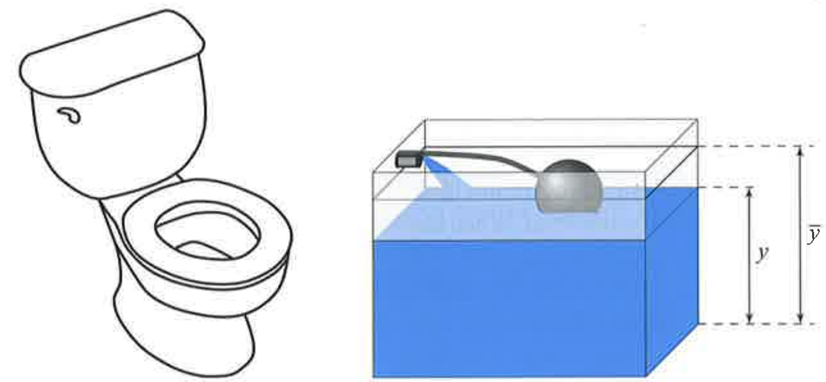
Compare the above analysis with the change in velocity produced by the open-loop solution (see P2.3):

$$\Delta y(t) = \left(1 - e^{-(b/m)t}\right) G(0)\bar{w}, \qquad G(0) = \frac{p}{b}. \tag{2.22}$$

Because for any $K > 0$ we have

$$G(0) = \frac{p}{b} > \frac{p}{b}\frac{1}{1 + (p/b)K} = D(0),$$

we conclude that the feedback solution always provides better *regulation* of the velocity in the presence of a road slope disturbance. Finally, large gains will bring down not only the tracking error but also the *regulation error* in response to a disturbance. Indeed, bigger Ks make both $S(0)$ and $D(0)$ small.



(a) Toilet with a water tank behind (b) Simplified toilet water tank with ballcock valve

Figure 2.15 A toilet with a water tank and a simplified schematic diagram showing the ballcock valve. The tank is in the shape of a rectangular prism with cross-sectional area A. The water level is y and the fill line is $\bar{y}$.

2.8 Integral Action

We close this chapter with an analysis of a simple and familiar controlled system: a toilet water tank, Fig. 2.15. This system has a property of much interest in control, the so-called *integral action*. As seen in previous sections, large gains will generally lead to small tracking errors but with potentially damaging consequences, such as large control inputs that can lead to saturation and other nonlinear effects. In the examples presented so far, only an infinitely large gain could provide *zero* steady-state tracking error in closed-loop. As we will see in this section, integral action will allow closed-loop systems to track constant references with zero steady-state tracking error without resorting to infinite gains.

A schematic diagram of a toilet water tank is shown in Fig. 2.15(b). Assuming that the tank has a constant cross-sectional area, A, the amount of water in the tank, i.e. the volume, v, is related to the water level, y, by

$$v = Ay.$$

When the tank is closed, for instance, right after a complete flush, water flows in at a rate $u(t)$, which is controlled by the ballcock valve. Without leaks, the water volume in the tank is preserved, hence

$$\dot{v}(t) = u(t).$$

On combining these two equations in terms of the water level, y, we obtain the differential equation

$$\dot{y}(t) = \frac{1}{A}u(t), \tag{2.23}$$

which reveals that the toilet water tank is essentially a *flow integrator*, as shown in the block-diagram representation in Fig. 2.16.

Figure 2.16 Block-diagram for water tank.

A *ballcock valve*, shown in Fig. 2.15(b), controls the inflow of water by using a float to measure the water level. When the water level reaches the *fill line*, $\bar{y}$, a lever connected to the float shuts down the valve. When the water level is below the fill line, such as right after a flush, the float descends and actuates the fill valve. This is a feedback mechanism. Indeed, we can express the flow valve as a function of the *error* between the fill line, $\bar{y}$, and the current water level, y, through

$$u(t) = K(\bar{y} - y),$$

where the profile of the function K is similar to the saturation curves encountered before in Figs. 1.4–1.6. The complete system is represented in the block-diagram Fig. 2.17, which shows that the valve is indeed a feedback element: the water level, y, tracks the reference level, fill line, $\bar{y}$.

With simplicity in mind, assume that the valve is linear. In this case, the behavior of the tank with the ballcock valve is given by the differential equation

$$\dot{y}(t) + \frac{K}{A}y(t) = \frac{K}{A}\bar{y}(t).$$

This equation is of the form (2.3) and has once again as solution (2.6), that is,

$$y(t) = \bar{y}(1 - e^{\lambda t}) + y_0 e^{\lambda t}, \quad \lambda = -\frac{K}{A}, \qquad t \geq 0.$$

Note, however, the remarkable fact that

$$\lim_{t \to \infty} y(t) = \bar{y}.$$

In other words, the steady-state solution is always equal to the target fill line, $\bar{y}$, if $K/A > 0$, no matter what the actual values of K and A are! The toilet water tank level, y, tracks the fill line level, $\bar{y}$, exactly without a high-gain feedback controller. As will become clear in Chapter 4, the reason for this remarkable property is the presence of a pure integrator in the feedback loop. Of course, the values of K and A do not affect the steady-state solution but do influence the rate at which the system converges to it.

Integral action can be understood with the help of the closed-loop diagram in Fig. 2.17. First note that for the output of an integrator to converge to a constant

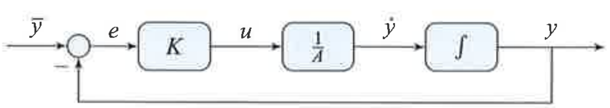


Figure 2.17 Block-diagram for water tank with ballcock valve.

value it is necessary that its input converge to zero. In Fig. 2.17, it is necessary[12] that $\lim_{t\to\infty} \dot{y}(t) = 0$. With that in mind, if $y(t)$ converges to anything other than $\bar{y}$, that is, $\lim_{t\to\infty} y(t) = \tilde{y} \neq \bar{y}$, then $\lim_{t\to\infty} e(t) = \tilde{y} - \bar{y} \neq 0$. But, if this is the case, $\lim_{t\to\infty}(A/K)\dot{y}(t) = \lim_{t\to\infty} e(t) = \tilde{y} - \bar{y} \neq 0$. Consequently $y(t)$ cannot converge to a constant other than $\bar{y}$. This is true even in the presence of some common nonlinearities in the loop. The ability to track constant references without high gains is the main reason behind the widespread use of integral control. We will analyze integral controllers in more detail in many parts of this book.

We conclude this section by revisiting the car example with linear model (2.3), by noting that when there is no damping, i.e. $b = 0$, then the car becomes a pure integrator. Indeed, in this case

$$\lim_{b\to 0} H(0) = \lim_{b\to 0} \frac{(p/b)K}{1 + (p/b)K} = 1, \qquad \lim_{b\to 0} S(0) = \lim_{b\to 0} \frac{1}{1 + (p/b)K} = 0,$$

which implies

$$\lim_{t\to\infty} y(t) = \bar{y} = \lim_{b\to 0} H(0)\bar{y} = \bar{y},$$

independently of the value of K. We saw in Section 2.7 that large controller gains lead not only to small tracking errors but also to effective disturbance rejection. The same is true for an integrator *in the controller*, which leads to asymptotic tracking *and* asymptotic disturbance rejection. However, the position of the integrator in the loop matters: an integrator in the system but not in the controller will lead to zero tracking error but nonzero disturbance rejection error. For instance, in the example of the car we have seen that $b \to 0$ implies $S(0) \to 0$ but

$$\lim_{b\to 0} D(0) = \lim_{b\to 0} \frac{p}{b}\frac{1}{1 + (p/b)K} = \frac{1}{K}, \qquad \lim_{t\to\infty} \Delta y(t) = D(0)\bar{w},$$

which is in general not zero. Nevertheless, it does gets smaller as the gain, K, gets large. By contrast, an integrator on the controller will generally lead to $S(0) = D(0) = 0$ independently of the loop gain, K. We will study this issue in more detail in Section 4.5.

Problems

2.1 Consider the solution (2.6) to the first-order ordinary differential equation (2.3) where the constant parameters m, b, and p are from the car velocity dynamic model developed in Section 2.1. Assign compatible units to the signals and constants in (2.3) and calculate the corresponding units of the parameter λ, from (2.7), and the time-constant τ, from (2.8).

[12] This is a necessary only condition, since $\dot{y}(t) = (1+t)^{-1}$ is such that $\lim_{t\to\infty} \dot{y}(t) = 0$ but $\lim_{t\to\infty} \int_0^t \dot{y}(\tau)d\tau = \ln(1+t) = \infty$.

2.2 Use the Taylor series expansion of the function $f(x) = \tan^{-1}(x)$ to justify the approximation

$$c\tan(\alpha^{-1}y) \approx by$$

when $b = \alpha^{-1}c$.

2.3 Calculate the dynamic response, $y(t)$, of the open-loop car velocity model (2.19) when

$$y_0 = \bar{y}, \qquad u(t) = G(0)^{-1}\bar{y}, \qquad w(t) = \bar{w}, \qquad t \geq 0,$$

and $G(0) = p/b$. Calculate the change in speed $\Delta y(t) = y(t) - y_0$ and compare your answer with (2.22).

The next problems involve the motion of particle systems using Newton's law.

2.4 Show that the first-order ordinary differential equation

$$m\dot{v} + bv = mg$$

is a simplified description of the motion of an object of mass m dropping vertically under constant gravitational acceleration, g, and linear air resistance, $-bv$.

2.5 The first-order ordinary differential equation obtained in P2.4 can be seen as a dynamic system where the output is the vertical velocity, v, and the input is the gravitational force, mg. Calculate the solution to this equation. Consider $m = 1$ kg, $b = 10$ kg/s, $g = 10\,\text{m/s}^2$. Sketch or use MATLAB to plot the response, $v(t)$, when $v(0) = 0$, $v(0) = 1$ m/s, or $v(0) = -1$ m/s.

2.6 Calculate the vertical position, $x(t)$, corresponding to the velocity, $v(t)$, computed in P2.5. How does the vertical position, $x(t)$, relate to the height measured from the ground, $h(t)$, of a free-falling object? Use the same data as in P2.5 and sketch or use MATLAB to plot the position $x(t)$ and the height $h(t)$, when $x(0) = 0$, $h(0) = 1$, and $v(0) = 0$, $v(0) = 1$ m/s, or $v(0) = -1$ m/s for 1 s.

2.7 The first-order nonlinear ordinary differential equation

$$m\dot{v} + bv^2 = mg, \qquad v > 0,$$

is a simplified description of the motion of an object of mass m dropping vertically under constant gravitational acceleration, g, and quadratic air resistance, bv^2. Verify that

$$v(t) = \frac{1 + \alpha\, e^{\lambda t}}{1 - \alpha\, e^{\lambda t}}\tilde{v}, \qquad \tilde{v} = \sqrt{\frac{mg}{b}}, \qquad \alpha = \frac{v(0) - \tilde{v}}{v(0) + \tilde{v}}, \qquad \lambda = -\frac{2b\tilde{v}}{m},$$

is a solution to this differential equation.

2.8 A sky diver weighing 70 kg reaches a constant vertical speed of 200 km/h during the free-fall phase of the dive and a vertical speed of 20 km/h after the parachute is opened. Approximate each phase of the fall by the ordinary differential equation obtained in P2.4 and estimate the resistance coefficients using the given information. Use $g = 10\,\text{m/s}^2$. What are the time-constants in each phase? At what time and distance from the ground should the parachute be opened if the landing speed is to be less than or equal to 29 km/h? If a dive starts at a height of 4 km with zero vertical velocity at the moment of the jump and the parachute is opened 60 s into the dive, how long is the diver airborne?

2.9 Redo P2.8 using the nonlinear model from P2.7.

The next problems involve the planar rotation of a rigid body. Such systems can be approximately modeled by the first-order ordinary differential equation:

$$J\dot{\omega} = \tau,$$

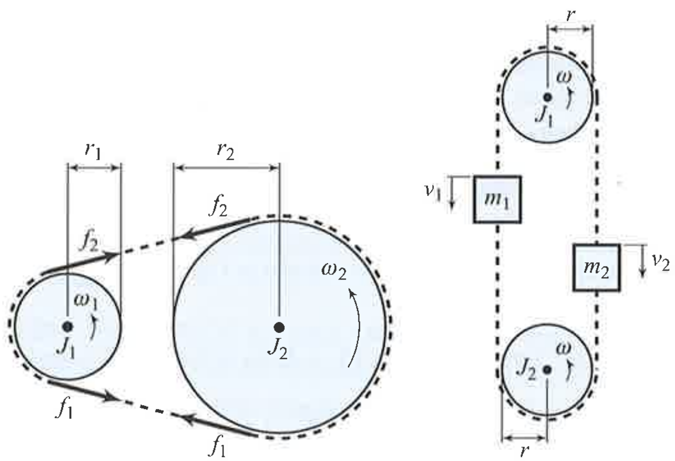
where ω is the body's angular speed, J is the body's moment of inertia about its center of mass, and τ is the sum of all torques about the center of mass of the body. In constrained rotational systems, e.g. lever, gears, etc., the center of mass can be replaced by the center of rotation.

2.10 An (inextensible and massless) belt is used to drive a rotating machine without slip as shown in Fig. 2.18(a). The simplified motion of the inertia J_1 is described by

$$J_1\,\dot{\omega}_1 = \tau + f_1 r_1 - f_2 r_1,$$

where τ is the torque applied by the driving motor and f_1 and f_2 are tensions on the belt. The machine is connected to the inertia J_2, which represents the sum of the inertias of all machine parts. The motion of the inertia J_2 is described by

$$J_2\,\dot{\omega}_2 = f_2 r_2 - f_1 r_2.$$



(a) Rotating machine (b) Elevator

Figure 2.18 Diagrams for P2.10 and P2.18.

Show that the motion of the entire system can be described by the differential equation

$$\left(J_1 r_2^2 + J_2 r_1^2\right) \dot{\omega}_1 = r_2^2 \tau, \qquad \omega_2 = (r_1/r_2)\,\omega_1.$$

2.11 Why is $f_1 \neq f_2$ in P2.10? Under what conditions are f_1 and f_2 equal?

2.12 Redo P2.10 in the presence of viscous friction torques, $-b_1\omega_1$ and $-b_2\omega_2$, on each inertia to obtain the differential equation

$$(J_1 r_2^2 + J_2 r_1^2)\dot{\omega}_1 + (b_1 r_2^2 + b_2 r_1^2)\omega_1 = r_2^2\,\tau.$$

2.13 Determine a first-order ordinary differential equation based on P2.10 and P2.12 to describe the rotating machine as a dynamic system where the output is the angular velocity of the inertia J_2, ω_2, and the input is the motor torque, τ. Calculate the solution to this equation. Consider $\tau = 1$ N m, $r_1 = 25$ mm, $r_2 = 500$ mm, $b_1 = 0.01$ kg m^2/s, $b_2 = 0.1$ kg m^2/s, $J_1 = 0.0031$ kg m^2, $J_2 = 25$ kg m^2. Sketch or use MATLAB to plot the response, $\omega_2(t)$, when $\omega_2(0) = 0$ rad/s, $\omega_2(0) = 3$ rad/s, or $\omega_2(0) = 6$ rad/s.

2.14 Calculate the (open-loop) motor torque, τ, for the rotating machine model in P2.10 and P2.12 so that the rotational speed of the mass J_2, ω_2, converges to $\bar{\omega}_2 = 4$ rad/s as t gets large. Use the same data as in P2.13 and sketch or use MATLAB to plot the response, $\omega_2(t)$, when $\omega_2(0) = 0$ rad/s, $\omega_2(0) = 3$ rad/s, or $\omega_2(0) = 6$ rad/s.

2.15 What happens with the response in P2.14 if the actual damping coefficients b_1 and b_2 are 20% larger than the ones you used to calculate the open-loop torque?

2.16 The feedback controller

$$\tau(t) = K(\bar{\omega}_2 - \omega_2(t))$$

can be used to control the speed of the inertia J_2 in the rotating machine discussed in P2.10 and P2.12. Calculate and solve a differential equation that describes the closed-loop response of the rotating machine. Using data from P2.13, select a controller gain, K, with which the time-constant of the rotating machine is 3 s. Compare your answer with the open-loop time-constant. Calculate the closed-loop steady-state error between the desired rotational speed, $\bar{\omega}_2 = 4$ rad/s, and $\omega_2(t)$. Sketch or use MATLAB to plot the response, $\omega_2(t)$, when $\omega_2(0) = 0$ rad/s, $\omega_2(0) = 3$ rad/s, or $\omega_2(0) = 6$ rad/s.

2.17 What happens with the response in P2.16 if the actual damping coefficients b_1 and b_2 are 20% larger than the ones you used to calculate the closed-loop gain?

2.18 A schematic diagram of an elevator is shown in Fig. 2.18(b). Proceed as in P2.10 to show that

$$\left(J_1 + J_2 + r^2(m_1 + m_2)\right)\dot{\omega} + (b_1 + b_2)\omega = \tau + gr(m_1 - m_2),$$

$v_1 = r\,\omega$, and $v_2 = -r\,\omega$, is a simplified description of the motion of the entire elevator system, where τ is the torque applied by the driving motor on the inertia J_1, and b_1 and b_2 are viscous friction torque coefficients at the inertias J_1 and J_2. If m_1 is the load to be lifted and m_2 is a counterweight, explain why is it advantageous to have the counterweight match the elevator load as closely as possible.

2.19 Determine a first-order ordinary differential equation based on P2.18 to describe the elevator as a dynamic system where the output is the vertical velocity of the mass m_1, v_1, and the inputs are the motor torque, τ, and the gravitational torque, $gr(m_1 - m_2)$. Calculate the solution to this equation. Consider $g = 10$ m/s^2, $\tau = 0$ N m, $r = 1$ m, $m_1 = m_2 = 1000$ kg, $b_1 = b_2 = 120$ kg m^2/s, $J_1 = J_2 = 20$ kg m^2. Sketch or use MATLAB to plot the response, $v_1(t)$, when $v_1(0) = 0$, $v_1(0) = 1$ m/s, or $v_1(0) = -1$ m/s.

2.20 Repeat P2.19 with $m_2 = 800$ kg.

2.21 Calculate the (open-loop) motor torque, τ, for the elevator model in P2.19 so that the vertical velocity of the mass m_1, v_1, converges to $\bar{v}_1 = 2$ m/s as t gets large. Use the same data as in P2.19 and sketch or use MATLAB to plot the response, $v_1(t)$, when $v_1(0) = 0$, $v_1(0) = 1$ m/s, or $v_1(0) = -1$ m/s.

2.22 Let $m_2 = 800$ kg and use the same motor torque, τ, you calculated in P2.21 and the rest of the data from P2.19 to sketch or use MATLAB to plot the response of the elevator mass m_1 velocity, $v_1(t)$, when $v_1(0) = 0$, $v_1(0) = 1$ m/s, or $v_1(0) = -1$ m/s. Did the velocity converge to $\bar{v}_1 = 2$ m/s? If not, recalculate a suitable torque. Plot the response with the modified torque, compare your answer with P2.21, and comment on the value of torque you obtained.

2.23 The feedback controller:

$$\tau(t) = K(\bar{v}_1 - v_1(t))$$

can be used to control the ascent and descent speed of the mass m_1 in the elevator discussed in P2.18 and P2.19. Calculate and solve a differential equation that describes the closed-loop response of the elevator. Using data from P2.19, select a controller gain, K, with which the time-constant of the elevator is approximately 5 s. Compare this value with the open-loop time-constant. Calculate the closed-loop steady-state error between the desired vertical velocity, $\bar{v}_1 = 2$ m/s, and $v_1(t)$. Sketch or use MATLAB to plot the response of the elevator mass m_1 velocity, $v_1(t)$, when $v_1(0) = 0$, $v_1(0) = 1$ m/s, or $v_1(0) = -1$ m/s. Compare the response with the open-loop control response from P2.22.

2.24 Repeat P2.23, this time setting the closed-loop time-constant to be about 0.5 s. What is the effect on the response? Do you see any problems with this solution?

2.25 Repeat P2.23 with $m_2 = 800$ kg. Treat the gravitational torque, $gr(m_1 - m_2)$, as a disturbance.

2.26 Repeat P2.25 this time setting the closed-loop time-constant to be about 0.5 s. What is the effect on the response? Do you see any problems with this solution?

Figures 2.19 through 2.21 show diagrams of mass–spring–damper systems. Assume that there is no friction between the wheels and the floor, and that all springs and dampers are linear: elongating a linear spring with rest length *ℓ_0 by $\Delta\ell$ produces an opposing force $k\Delta\ell$* (Hooke's Law)*, where $k > 0$ is the spring* stiffness*; changing the length of* *a linear damper at a rate $\dot{\ell}$ produces an opposing force $b\dot{\ell}$, where b is the damper's* damping coefficient.

2.27 Choose x_0 wisely to show that the ordinary differential equation

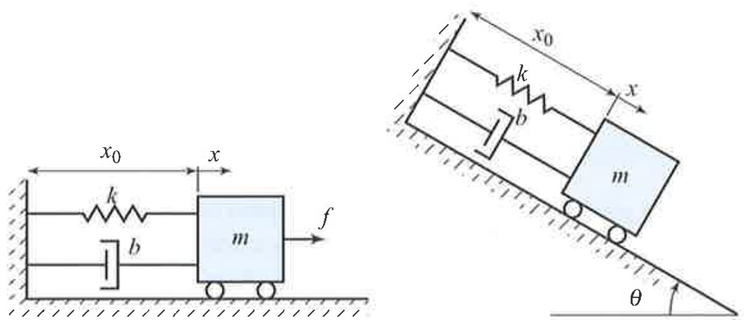
$$m\ddot{x} + b\dot{x} + kx = f$$

is a simplified description of the motion of the mass–spring–damper system in Fig. 2.19(a), where f is a force applied on mass m. Why does the equation not depend on the spring rest length ℓ_0?

2.28 Show that the ordinary differential equation

$$m\ddot{x} + b\dot{x} + kx = mg\sin\theta$$

is a simplified description of the motion of the mass–spring–damper system in Fig. 2.19(b), where g is the gravitational acceleration and x_0 is equal to the spring rest length ℓ_0.



(a) (b)

Figure 2.19 Diagrams for P2.27 and P2.28.

2.29 Rewrite the ordinary differential equation obtained in P2.28 as

$$m\ddot{y} + b\dot{y} + ky = 0, \qquad y = x - k^{-1}mg\sin\theta.$$

Relate this result to a different choice of x_0 and comment on your findings.

2.30 Show that the ordinary differential equation

$$m\ddot{x} + b\dot{x} + (k_1 + k_2)x = 0$$

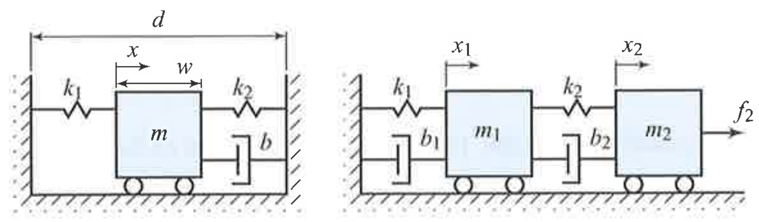
is a simplified description of the motion of the mass–spring–damper system in Fig. 2.20(a). What does x represent? Why do the equations not depend either on the rest lengths $\ell_{0,1}$, $\ell_{0,2}$ or on the dimensions d and w? What is the difference between the cases $d \geq w + \ell_{0,1} + \ell_{0,2}$ and $d \leq w + \ell_{0,1} + \ell_{0,2}$?

2.31 Can you replace the two springs in P2.30 by a single spring and still obtain the same ordinary differential equation?

2.32 Show that the ordinary differential equations

$$m_1\ddot{x}_1 + (b_1 + b_2)\dot{x}_1 + (k_1 + k_2)x_1 - b_2\dot{x}_2 - k_2x_2 = 0,$$
$$m_2\ddot{x}_2 + b_2(\dot{x}_2 - \dot{x}_1) + k_2(x_2 - x_1) = f_2$$

constitute a simplified description of the motion of the mass–spring–damper system in Fig. 2.20(b), where $x_1 = x_2 = 0$ when the length of both springs is equal to their rest lengths and f_2 is a force applied on mass m_2.



(a) (b)

Figure 2.20 Diagrams for P2.30 and P2.32.

2.33 Show that the ordinary differential equations

$$m_1\ddot{x}_1 + b(\dot{x}_1 - \dot{x}_2) + k(x_1 - x_2) = 0,$$
$$m_2\ddot{x}_2 + b(\dot{x}_2 - \dot{x}_1) + k(x_2 - x_1) = 0,$$

constitute a simplified description of the motion of the mass–spring–damper system in Fig. 2.21 where $x_1 = x_2 = 0$ when the length of the spring is its rest length. Show that it is possible to *decouple* the equations if you write them in the coordinates

$$y_1 = \frac{m_1x_1 + m_2x_2}{m_1 + m_2}, \qquad y_2 = x_1 - x_2.$$

Use what you know from physics to explain why.

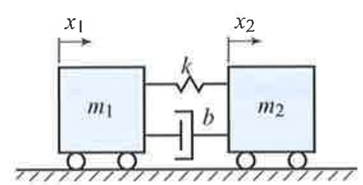


Figure 2.21 Diagram for P2.33.

The next problems have simple electric circuits. Electric circuits can be accurately modeled using ordinary differential equations.

2.34 An electric circuit in which a capacitor is in series with a resistor is shown in Fig. 2.22(a). In an electric circuit, the sum of the voltages around a loop must equal

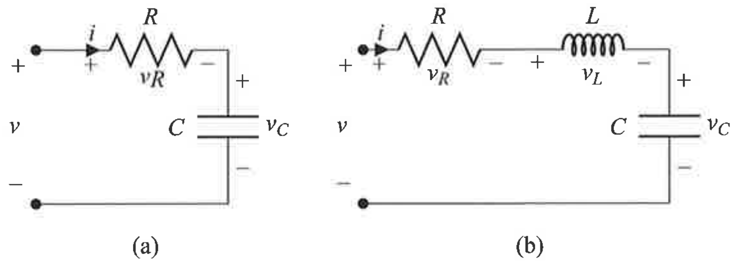


Figure 2.22 Diagrams for P2.34 and P2.36.

zero:

$$-v + v_R + v_C = 0.$$

This is Kirchhoff's voltage law. The voltage and the current on the capacitor and resistor satisfy

$$i_C = C\,\dot{v}_C, \qquad v_R = R\,i_R,$$

where C is the capacitor's *capacitance* and R is the resistor's *resistance*. In this circuit

$$i_R = i_C = i,$$

because all elements are *in series*. This is Kirchhoff's current law. Show that

$$R\,C\,\dot{v}_C + v_C = v$$

is the equation governing this *RC*-circuit.

2.35 Consider the *RC*-circuit from P2.34 where $R = 1\,\text{M}\Omega$ and $C = 10\,\mu\text{F}$. Assuming zero initial conditions, sketch the capacitor's voltage, $v_C(t)$, when a constant voltage $v(t) = 10\,\text{V}$, $t \geq 0$, is applied to the circuit. Sketch also the circuit current $i(t)$.

2.36 An electric circuit in which an inductor, a capacitor, and a resistor are in series is shown in Fig. 2.22(b). As in P2.34, the sum of the voltages around a loop must equal zero:

$$-v + v_R + v_L + v_C = 0.$$

This is Kirchhoff's voltage law. The voltages and the currents on the capacitor and resistor are as in P2.34 and the voltage on the inductor is

$$v_L = L\,\dot{i}_L,$$

where L is the inductor's *inductance*. Because the elements are in series

$$i_R = i_C = i_L = i.$$

This is Kirchhoff's current law. Show that

$$L\,C\,\ddot{v}_c + R\,C\,\dot{v}_C + v_C = v$$

is the equation governing the *RLC*-circuit.

2.37 Consider the differential equation for the *RLC*-circuit from P2.36. Compare this equation with the equations of the mass–spring–damper system from P2.27 and explain how one could select values of the resistance, R, capacitance, C, inductance, L, and input voltage, v, to simulate the movement of the mass–spring–damper system in P2.27. The resulting device is an *analog computer*.

2.38 An approximate model for the electric circuit in Fig. 2.23, where the triangular element is an amplifier with a very large gain (operational amplifier, OpAmp), is obtained from

$$R_1 i_{R_1} = v - v_-, \qquad i_{C_1} = C_1(\dot{v} - \dot{v}_-), \qquad i_{C_2} = C_2(\dot{v}_- - \dot{v}_o),$$

and

$$v_- \approx v_+ = 0, \qquad i_{C_2} = i_{C_1} + i_{R_1}.$$

Show that

$$R_1\,C_2\,\dot{v}_o + R_1\,C_1\,\dot{v} + v = 0.$$

Solve the auxiliary differential equation:

$$\dot{z} + \frac{1}{R_1\,C_2} v = 0$$

and show that

$$v_o(t) = R_1\,C_1\,\dot{z}(t) + z(t)$$

solves the original differential equation.

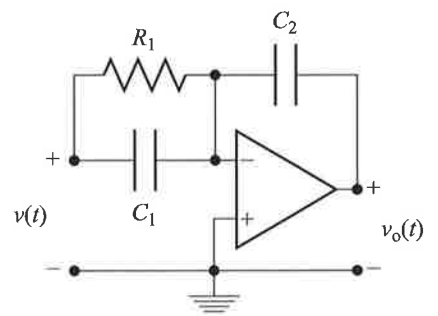


Figure 2.23 Diagram for P2.38.

2.39 Consider the OpAmp-circuit from P2.38 where $R_1 = 1\,\text{M}\Omega$ and $C_1 = C_2 = 10\,\mu\text{F}$. Assuming zero initial conditions, sketch the output voltage, $v_o(t)$, when a constant voltage $v(t) = 10\,\text{V}$, $t \geq 0$, is applied to the circuit.

2.40 In P2.38, set $C_1 = 0$ and solve for $v_o(t)$ in terms of $v(t)$. Name one application for this circuit.

2.41 The mechanical motion of the rotor of a DC motor shown schematically in Fig. 2.24(a) can be described by the differential equation

$$J\,\dot{\omega} + b\,\omega = \tau,$$

where ω is the rotor angular speed, J is the rotor moment of inertia, b is the coefficient of viscous friction. The rotor torque, τ, is given by

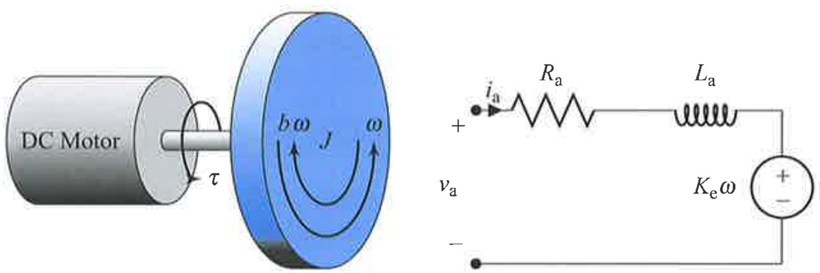
$$\tau = K_t\, i_a$$

where i_a is the armature current and K_t is the motor torque constant. Neglecting the effects of the armature inductance ($L_a \approx 0$), the current is determined by the circuit in Fig. 2.24(b):

$$v_a = R_a\, i_a + K_e\,\omega,$$

where v_a is the armature voltage, R_a is the armature resistance, and K_e is the back-EMF constant. Combine these equations to show that

$$J\,\dot{\omega} + \left(b + \frac{K_e K_t}{R_a}\right)\omega = \frac{K_t}{R_a} v_a.$$



(a) Mechanical (b) Electric

Figure 2.24 Diagrams for P2.41.

2.42 Show that $K_t = K_e$. *Hint: Equate the mechanical power with the electric power.*

2.43 The first-order ordinary differential equation obtained in P2.41 can be seen as a dynamic system where the output is the angular velocity, ω, and the input is the armature voltage, v_a. Calculate the solution to this equation when v_a is constant. Estimate the parameters of the first-order differential equation describing a DC motor that achieves a steady-state angular velocity of 5000 RPM when $v_a = 12$ V and has a time-constant of 0.1 s. Can you also estimate the "physical" parameters J, b, K_t, K_e, and R_a with this information?

2.44 Can you estimate the parameters J, K_t, K_e, and b of the DC motor in P2.43 if you know $R_a = 0.2\,\Omega$ and the *stall torque* $\tau = 1.2$ N m at $v_a = 12$ V? *Hint: The stall torque is attained when the motor is held in place.*

2.45 DC motors with high-ratio gear boxes can be damaged if held in place. Can you estimate the parameters J, K_t, K_e, and b of the DC motor in P2.43 if you know that $R_a = 0.2\,\Omega$ and that after you attach an additional inertia $J' = 0.001$ kg m^2 the motor time-constant becomes 0.54 s?

2.46 Redo P2.41 using the equations in P2.36 to show that

$$J L_a\,\ddot{\omega} + (J R_a + b L_a)\dot{\omega} + (K_e K_t + b R_a)\,\omega = K_t\, v_a$$

when $L_a > 0$ is not negligible. Show that this equation reduces to the one in P2.41 if $L_a = 0$.

2.47 The feedback controller

$$v_a(t) = K(\bar{\omega} - \omega(t))$$

can be used to regulate the angular speed of the DC motor, $\omega(t)$, for which a model was developed in P2.41. Calculate and solve a differential equation that describes the closed-loop response of the DC motor. Using data from P2.43, select a controller gain, K, with which the closed-loop steady-state error between the desired angular speed, $\bar{\omega}$, and the actual angular speed, $\omega(t)$, is less than 10%. Calculate the resulting closed-loop time-constant and sketch or use MATLAB to plot the output $\omega(t)$ and the voltage $v_a(t)$ generated in response to a reference $\bar{\omega} = 4000$ RPM assuming zero initial conditions. What is the maximum value of $v_a(t)$?

2.48 Redo P2.47 but this time design K such that $v_a(0)$ is always smaller than 12 V when $\bar{\omega} = 4000$ RPM.

The next problems have simple examples of heat and fluid flow using ordinary differential equations. Detailed modeling of such phenomena often requires partial differential equations.

2.49 The temperature, T (in K or in °C), of a substance flowing in and out of a container kept at the ambient temperature, T_o, with an inflow temperature, T_i, and a heat source, q (in W), can be approximated by the differential equation

$$mc\,\dot{T} = q + wc\,(T_i - T) + \frac{1}{R}(T_o - T),$$

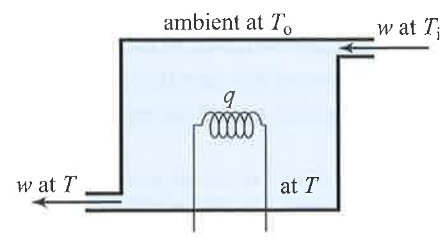


Figure 2.25 Diagram for P2.49.

where m and c are the substance's mass and specific heat, and R is the overall system's thermal resistance. The input and output flow mass rates are assumed to be equal to w. This differential equation model can be seen as a dynamic system where the output is the substance's temperature, T, and the inputs are the heat source, q, the flow rate, w, and the temperatures T_o and T_i. Calculate the solution to this equation when q, w, T_o, and T_i are constants.

2.50 Assume that water's density and specific heat are $997.1\,\text{kg/m}^3$ and $c = 4186\,\text{J/kg K}$. A 50 gal ($\approx 0.19\,\text{m}^3$) water heater is turned off full with water at 140 °F ($\approx$ 60 °C). Use the differential equation in P2.49 to estimate the heater's thermal resistance, R, knowing that after 7 days left at a constant ambient temperature, 77 °F ($\approx$ 25 °C), without turning it on, $q = 0$, or cycling any water, $w = 0$, the temperature of the water was about 80 °F ($\approx$ 27 °C).

2.51 For the same conditions as in P2.50, calculate how much time it takes for the water temperature to reach 80 °F ($\approx$ 27 °C) with a constant in/out flow of 20 gal/h ($\approx 21 \times 10^{-6}\,\text{m}^3/\text{s}$) at ambient temperature. Compare your answer with the case when no water flows through the water heater.

2.52 Consider a water heater as in P2.50 rated at 40,000 BTU/h ($\approx$ 12 kW). Calculate the time it takes to heat up a heater initially full with water at ambient temperature to 140 °F ($\approx$ 60 °C) without any in/out flow of water, $w = 0$.

2.53 Repeat P2.52 for a constant in/out flow of 20 gal/h at ambient temperature. Compare the solutions.

2.54 Most residential water heaters have a simple *on/off*-type controller: the water heater is turned on at full power when the water temperature, T, falls below a set value, $\underline{T}$, and is turned off when it reaches a second set point, $\overline{T}$. For a 50 gal ($\approx 0.19\,\text{m}^3$) heater as in P2.50 rated at 40,000 BTU/h ($\approx$ 12 kW) and with thermal resistance $R = 0.27$ K/W, sketch or use MATLAB to plot the temperature of the water during 24 hours for a heater with an on/off controller set with $\underline{T} = 122$ °F ($\approx$ 50 °C) and $\overline{T} = 140$ °F ($\approx$ 60 °C), without any in/out flow of water, $w = 0$. Assume that the heater is initially full with water a tad below $\underline{T}$. Compute the average water temperature and power consumption for a complete on/off cycle.

2.55 Repeat P2.54 for a constant in/out flow of 20 gal/h at ambient temperature. Compare the solutions.

2.56 Repeat P2.54 with $\underline{T} = 129.2$ °F ($\approx$ 54 °C) and $\overline{T} = 132.8$ °F ($\approx$ 56 °C). What is the impact of the choice of $\underline{T}$ and $\overline{T}$ on the performance of the controller?

3 Transfer-Function Models

The dynamic models developed in Chapter 2 relate input and output signals evolving in the *time domain* through differential equations. In this chapter we will introduce transform methods that can relate input and output signals in the *frequency domain*, establishing a correspondence between a *time-invariant* linear system model and its frequency-domain *transfer-function*. A linear system model may not be time-invariant but, for linear time-invariant systems, the frequency domain provides an alternative vantage point from which to perform calculations and interpret the behavior of the system and the associated signals. This perspective will be essential to many of the control design methods to be introduced later in this book, especially those of Chapter 7.

Frequency-domain models can be obtained experimentally or derived formally from differential equations using the Laplace or Fourier transforms. In controls we work mostly with the Laplace transform. This chapter starts with a brief but not completely elementary review of the Laplace transform before returning to signals, systems, and controls. From this chapter on, transfer-functions become functions of complex variables and we use the symbol j to represents the imaginary unit, that is $j = \sqrt{-1}$. We use $\dot{f}(t)$ and $f'(s)$ to distinguish between differentiation with respect to the real variable $t \in \mathbb{R}$ or differentiation with respect to the complex variable $s \in \mathbb{C}$.

3.1 The Laplace Transform

The Laplace transform is essentially a sophisticated *change-of-variables* that associates a function[1] $f(t)$ of the *real-valued* time variable $t \in \mathbb{R}$, $t \geq 0$, with a function $F(s)$ of a *complex-valued* frequency variable $s \in \mathbb{C}$. Formally, the Laplace transform of the function $f(t)$ is the result of the integral:[2]

$$F(s) = \mathscr{L}\{f(t)\} = \int_{0^-}^{\infty} f(t)e^{-st}\,dt. \tag{3.1}$$

This integral may not converge for every function $f(t)$ or every s, and when it converges it may not have a *closed-form* solution. However, for a large number of common

[1] A formal setup that is comfortable is that of piecewise continuous or piecewise smooth (continuous and infinitely differentiable) functions with only a discrete set of discontinuities, such as the one adopted in [LeP10].

[2] The notation 0^- means the one-side limit $\lim_{\epsilon \uparrow 0} \epsilon$, which is used to accommodate possible discontinuities of $f(t)$ at the origin.