4

Simple Control Systems

4.1 Introduction

In this chapter we will give simple examples of analysis and design of control systems. We will start in Sections 4.2 and 4.3 with two systems that can be handled using only knowledge of differential equations. Section 4.2 deals with design of a cruise controller for a car. In Section 4.3 we discuss the dynamics of a bicycle, many of its nice properties are due to a purely mechanical feedback which has emerged as a result of trial and error over a long period of time. Section 3.3 is a suitable preparation for Sections 4.2 and 4.3. Differential equations are cumbersome for more complicated problems and better tools are needed. Efficient methods for working with linear systems can be developed based on a basic knowledge of Laplace transforms and transfer functions. Coupled with block diagrams this gives a very efficient way to deal with linear systems. The block diagram gives the overview and the behavior of the individual blocks are described by transfer functions. The Laplace transforms make it easy to manipulate the system formally and to derive relations between different signals. This is one of the standard methods for working with control systems. It is exploited in Section 4.4, which gives a systematic way of designing PI controllers for first order systems. This section also contains material required to develop an intuitive picture of the properties of second order systems. Section 4.5 deals with design of PI and PID controllers for second order systems. A proper background for Sections 4.4 and 4.5 is Section 3.4. Section 4.6 deals with the design problem for systems of arbitrary order. This section which requires more mathematical maturity can be omitted in a first reading. For the interested reader it gives, however, important insight into the design problem and the structure of stabilizing controllers. Section 4.6 summarizes the chapter and

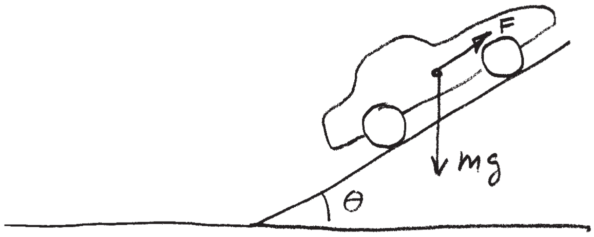

Figure 4.1 Schematic diagram of a car on a sloping road.

outlines some important issues that should be considered.

4.2 Cruise Control

The purpose of cruise control is to keep the velocity of a car constant. The driver drives the car at the desired speed, the cruise control system is activated by pushing a button and the system then keeps the speed constant. The major disturbance comes from changes of the slope of the road which generates forces on the car due to gravity. There are also disturbances due to air and rolling resistance. The cruise control system measures the difference between the desired and the actual velocity and generates a feedback signal which attempts to keep the error small in spite of changes in the slope of the road. The feedback signal is sent to an actuator which influences the throttle and thus the force generated by the engine.

We will start by developing a mathematical model of the system. The mathematical model should tell how the velocity of the car is influenced by the throttle and the slope of the road. A schematic picture is shown in Figure 4.1

Modeling

We will model the system by a momentum balance. The major part of the momentum is the product of the velocity v and the mass m of the car. There are also momenta stored in the engine, in terms of the rotation of the crank shaft and the velocities of the cylinders, but these are much smaller than mv. Let θ denote the slope of the road, the momentum balance can be

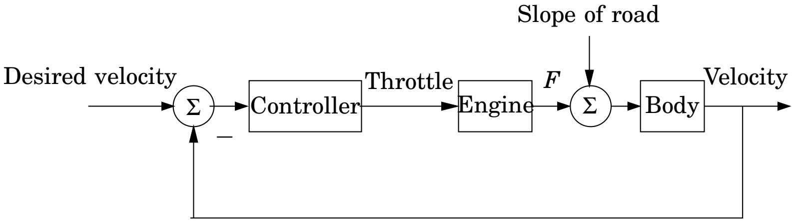

Figure 4.2 Block diagram of a car with cruise control.

written as

$$m\frac{dv}{dt} + cv = F - mg\theta \tag{4.1}$$

where the term cv describes the momentum loss due to air resistance and rolling and F is the force generated by the engine. The retarding force due to the slope of the road should similarly be proportional to the sine of the angle but we have approximated $\sin\theta \approx \theta$. The consequence of the approximations will be discussed later. It is also assumed that the force F developed by the engine is proportional to the signal u sent to the throttle. Introducing parameters for a particular car, an Audi in fourth gear, the model becomes

$$\frac{dv}{dt} + 0.02v = u - 10\theta \tag{4.2}$$

where the control signal is normalized to be in the interval $0 \le u \le 1$, where $u = 1$ corresponds to full throttle. The model implies that with full throttle in fourth gear the car cannot climb a road that is steeper than 10%, and that the maximum speed in 4th gear on a horizontal road is $v = 1/0.02 = 50$ m/s (180 km/hour).

Since it is desirable that the controller should be able to maintain constant speed during stationary conditions it is natural to choose a controller with integral action. A PI controller is a reasonable choice. Such a controller can be described by

$$u = k(v_r - v) + k_i \int_0^t (v_r - v(\tau))d\tau \tag{4.3}$$

A block diagram of the system is shown in Figure 4.2. To understand how the cruise control system works we will derive the equations for the closed loop systems described by Equations (4.2) and (4.3) Since the effect of the slope on the velocity is of primary interest we will derive an equation that tells how the velocity error $e = v_r - v$ depends on the slope of the road.

Assuming that v_r is constant we find that

$$\frac{dv}{dt} = -\frac{de}{dt}, \quad \frac{d^2v}{dt^2} = -\frac{d^2e}{dt^2}$$

It is convenient to differentiate (4.3) to avoid dealing both with integrals and derivatives. Differentiating the process model (4.2) the term du/dt can be eliminated and we find the following equation that describes the closed loop system

$$\frac{d^2e}{dt^2} + (0.02 + k)\frac{de}{dt} + k_i e = 10\frac{d\theta}{dt} \tag{4.4}$$

We can first observe that if θ and e are constant the error is zero. This is no surprise since the controller has integral action, see the discussion about the integral action Section 2.2.

To understand the effects of the controller parameters k and k_i we can make an analogy between (4.4) and the differential equation for a mass-spring-damper system

$$M\frac{d^2x}{dt^2} + D\frac{dx}{dt} + Kx = 0$$

We can thus conclude that parameter k influences damping and that k_i influences stiffness.

The closed loop system (4.4) is of second order and it has the characteristic polynomial

$$s^2 + (0.02 + k)s + k_i \tag{4.5}$$

We can immediately conclude that the roots of this polynomial can be given arbitrary values by choosing the controller parameters properly. To find reasonable values we compare the characteristic polynomial with the characteristic polynomial of the normalized second order polynomial

$$s^2 + 2\zeta\omega_0 s + \omega_0^2 \tag{4.6}$$

where ζ denotes relative damping and ω_0 is the undamped natural frequency. The parameter ω_0 gives response speed, and ζ determines the shape of the response. Comparing the coefficients of the closed loop characteristic polynomial (4.5) with the standard second order polynomial (4.6) we find that the controller parameters are given by

$$\begin{aligned} k &= 2\zeta\omega_0 - 0.02 \\ k_i &= \omega_0^2 \end{aligned} \tag{4.7}$$

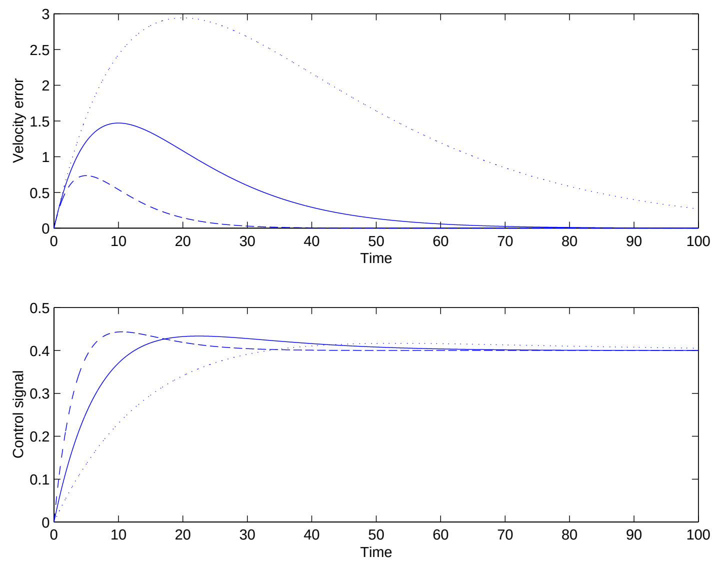

Figure 4.3 Simulation of a car with cruise control for a step change in the slope of the road. The controllers are designed with relative damping $\zeta = 1$ and $\omega_0 = 0.05$ (dotted), $\omega_0 = 0.1$ (full) and $\omega_0 = 0.2$ (dashed).

Since it is desirable that a cruise control system should respond to changes in the slope in a smooth manner without oscillations it is natural to choose $\zeta = 1$, which corresponds to critical damping. Then there is only one parameter ω_0 that has to be determined. The selection of this parameter is a compromise between response speed and control actions. This is illustrated in Figure 4.3 which shows the velocity error and the control signal for a simulation where the slope of the road suddenly changes by 4%. Notice that the largest velocity error decreases with increasing ω_0, but also that the control signal increases more rapidly. In the simple model (4.1) it was assumed that the force responded instantaneously to the throttle. For rapid changes there may be additional dynamics that has to be accounted for. There are also physical limitations to the rate of change of the force. These limitations, which are not accounted for in the simple model (4.1), limit the admissible value of ω_0. Figure 4.3 shows the velocity error and the control signal for a few values of ω_0. A reasonable choice of ω_0 is in the range of 0.1 to 0.2. The performance of the cruise control system can be evaluated by comparing the behaviors of cars with and without cruise control. This is done in Figure 4.4 which shows the velocity error when the slope of the road is suddenly increased by 4%. Notice the drastic

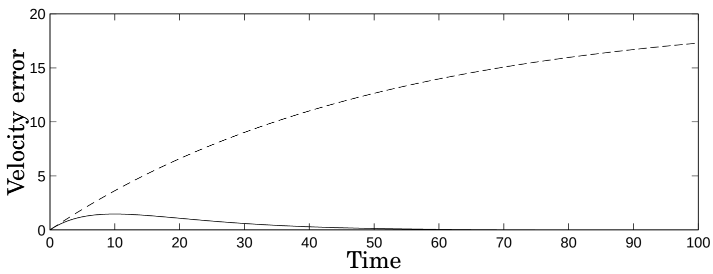

Figure 4.4 Simulation of a car with (solid line) and without cruise control (dashed line) for a step change of 4% in the slope of the road. The controller is designed for $\omega_0 = 0.1$ and $\zeta = 1$.

difference between the open and closed loop systems.

With the chosen parameters $\omega_0 = 0.2$ and $\zeta = 1$ we have $2\zeta\omega_0 = 0.2$ and it follows from (4.7) that the parameter $c = 0.02$ has little influence on the behavior of the closed loop system since it is an order of magnitude smaller than $2\zeta\omega_0$. Therefore it is not necessary to have a very precise value of this parameter. This is an illustration of an important and surprising property of feedback, namely that feedback systems can be designed based on simplified models. This will be discussed extensively in Chapter 5.

A cruise control system contains much more than an implementation of the PI controller given by (4.3). The human-machine interface is particularly important because the driver must be able to activate and deactivate the system and to change the desired velocity. There is also logic for deactivating the system when braking, accelerating or shifting gear.

4.3 Bicycle Dynamics

The bicycle is an ingenious device for recreation and transportation, which has evolved over a long period of time. It is a very effective vehicle that is extremely maneverable. Feedback is essential for understanding how the bicycle really works. In the bicycle there is no explicit control system with sensing and actuation, instead control is accomplished by clever mechanical design of the front fork which creates a feedback that under certain conditions stabilizes the bicycle. It is worth mentioning that the literature on bicycles is full of mistakes or misleading statements. We quote from the book Bicycling Science by Whitt and Wilson:

> The scientific literature (Timoshenko, Young, DenHartog et. al.) shows

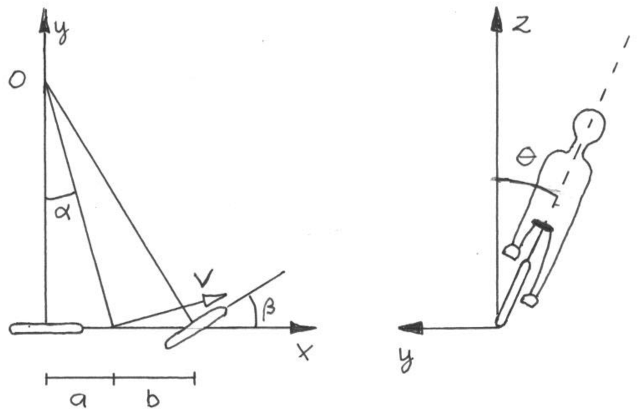

Figure 4.5 Schematic picture of a bicycle. The top view is shown on the left and the rear view on the right.

> often complete disagreement even about fundamentals. One advocates that a high center of mass improves stability, another concludes that a low center of mass is desirable.

We start by developing a simple modeling that clearly shows that feedback is an essential aspect of a bicycle.

Modeling

A detailed model of the bicycle is quite complicated. We will derive a simplified model that captures many relevant balancing properties of the bicycle. To understand how a bicycle works it is necessary to consider the system consisting of the bicycle and the rider. The rider can influence the bicycle in two ways by exerting a torque on the handle bar and by leaning. We will neglect the lean and consider the rider as a rigid body, firmly attached to the bicycle. A schematic picture of the bicycle is shown in Figure 4.5. To describe the dynamics we must account for the tilt of the bicycle. We introduce a coordinate system fixed to the bicycle with the x-axis through the contact points of the wheels with the ground, the y-axis horizontal and the z-axis vertical, as shown in Figure 4.5. Let m be the mass of the bicycle and the rider, J the moment of inertia of the bicycle and the rider with respect to the x-axis. Furthermore let l be the distance from the x-axis to the center of mass of bicycle and rider, θ the tilt angle and F the component of the force acting on rider and the bicycle.

A momentum balance around the x-axis gives

$$J\frac{d^2\theta}{dt^2} = mgl \sin\theta + Fl \cos\theta \tag{4.8}$$

The force F has two components, a centripetal force and an inertia force due to the acceleration of the coordinate system. The force can be determined from kinematic relations, see Figure 4.5. To describe these we introduce the steering angle β, and the forward velocity V_0. Furthermore the distance between the contact point of the front and rear wheel is b and the distance between the contact point of the rear wheel and the projection of the center of mass of bicycle and rider is a. To simplify the equations it is assumed that the angles β and θ are so small that sines and tangent are equal to the angle and cosine is equal to the one. Viewed from the top as shown in Figure 4.5 the bicycle has its center of rotation at a distance b/θ from the rear wheel. The centripetal force is

$$F_c = \frac{mV_0^2}{b}\beta$$

The y-component of the velocity of the center of mass is

$$V_y = V_0\alpha = \frac{aV_0}{b}\beta$$

where a is the distance from the contact point of the back wheel to the projection of the center of mass. The inertial force due to the acceleration of the coordinate system is thus

$$F_i = \frac{amV_0}{b}\frac{d\beta}{dt}$$

Inserting the total force $F = F_c + F_i$ into (4.8) we find that the bicycle can be described by

$$J\frac{d^2\theta}{dt^2} = mgl\theta + \frac{amV_0 l}{b}\frac{d\beta}{dt} + \frac{mV_0^2 l}{b}\beta \tag{4.9}$$

This equation has the characteristic equation

$$Js^2 - mgl = 0$$

which has the roots

$$s = \pm\sqrt{\frac{mgl}{J}}$$

The system is unstable, because the characteristic equation has one root in the right half plane. We may therefore believe that the rider must actively stabilize the bicycle all the time.

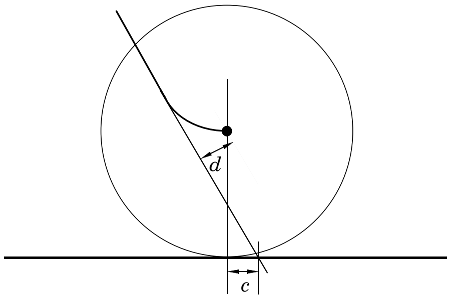

Figure 4.6 Schematic picture of the front fork.

The Front Fork

The bicycle has a front fork of rather intriguing design, see Figure 4.6. The front fork is angled and shaped so that the contact point of the wheel with the road is behind the axis of rotation of the front wheel assembly. The distance c is called the trail. The effect of this is that there will be a torque on the front wheel assembly when the bicycle is tilted. Because of the elasticity of the wheel there will be a compliance that also exerts a torque. The driver will also exert a torque on the front wheel assembly. Let T be the torque applied on the front fork by the driver. A static torque balance for the front fork assembly gives

$$\beta = -k_1\theta + k_2 T \tag{4.10}$$

Strictly speaking we should have a differential equation, for simplicity we will use the static equation.

Taking the action of the front fork into account we find that the bicycle is described by the Equations 4.9 and 4.10. A block diagram of representation of the system is shown in Figure 4.7. The figure shows clearly that the bicycle with the front fork is a feedback system. The front wheel angle β influences the tilt angle θ as described by (4.9) and the tilt angle influences the front wheel angle as described by (4.10). We will now investigate the consequences of the feedback created by the front fork. Inserting the expression (4.10) for steering angle β in the momentum balance (4.9) we get

$$J\frac{d^2\theta}{dt^2} + \frac{amV_0lk_1}{b}\frac{d\theta}{dt} + \Big(\frac{mV_0^2lk_1}{b} - mgl\Big)\theta = \frac{amV_0lk_2}{b}\frac{dT}{dt} + \frac{mV_0^2k_2l}{b}T \tag{4.11}$$

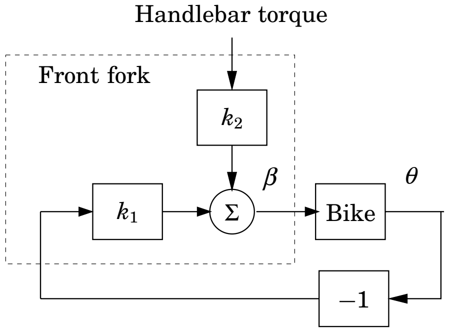

Figure 4.7 Block diagram of a bicycle with the front fork.

The characteristic equation of this system is

$$Js^2 + \frac{amV_0lk_1}{b}s + \Big(\frac{mV_0^2lk_1}{b} - mgl\Big) = 0$$

This equation is stable if

$$V_0 > V_c = \sqrt{\frac{gb}{k_1}} \qquad (4.12)$$

We can thus conclude that because of the feedback created by the design of the front fork the bicycle will be stable provided that the velocity is sufficiently large. The velocity V_c is called the critical velocity.

Useful information about bicycle dynamics can be obtained by driving it with constant speed V_0 in a circle with radius r_0. To determine the numerical values of the essential parameters a torque wrench can be used to measure the torque the driver exerts on the handle bar. In steady state conditions the centripetal force must be balanced by the gravity. Assuming that the bicycle moves counter clockwise the lean angle is

$$\theta_0 = -\frac{V_0^2}{r_0g}$$

It then follows from (4.11) that the torque required is given by

$$T_0 = \frac{bgl - V_0^2lk_1)}{k_2lr_0g} = \frac{k_1(V_c^2 - V_0^2)}{k_2r_0g}$$

This means that no torque is required if the bicycle is driven at the critical velocity and that the torque changes sign at the critical velocity.

Rear-wheel Steering

The analysis performed shows that feedback analysis gives substantial insight into behavior of bicycles. Feedback analysis can also indicate that a proposed system may have substantial disadvantages that are not apparent from static analysis. It is therefore essential to consider feedback and dynamics at an early stage of design. We illustrate this with a bicycle example. There are advantages in having rear-wheel steering on recumbent bicycles because the design of the drive is simpler. Again we quote from Whitt and Wilson Bicycling Science:

> The U.S. Department of Transportation commissioned the construction of a safe motorcycle with this configuration (rear-wheel steering). It turned out to be safe in an unexpected way: No one could ride it.

The reason for this is that a bicycle with rear-wheel steering has dynamics which makes it very difficult to ride. This will be discussed in Sections 5.9. Let it suffice to mention that it is essential to consider dynamics and control at an early stage of the design process. This is probable the most important reason why all engineers should have a basic knowledge about control.

4.4 Control of First Order Systems

We will now develop a systematic procedure for finding controllers for simple systems. To do this we will be using the formalism based on Laplace transforms and transfer functions which is developed in Section 3.4. This simplifies the calculations required substantially. In this section we will consider systems whose dynamics are of first order differential equations. Many systems can be approximately described by such equations. The approximation is reasonable if the storage of mass, momentum and energy can be captured by one state variable. Typical examples are

- Velocity of car on the road
- Control of velocity of rotating system
- Electric systems where energy is essentially stored in one component
- Incompressible fluid flow in a pipe
- Level control of a tank
- Pressure control in a gas tank
- Temperature in a body with essentially uniform temperature distribution e.g., a vessel filled with a mixture of steam and water.

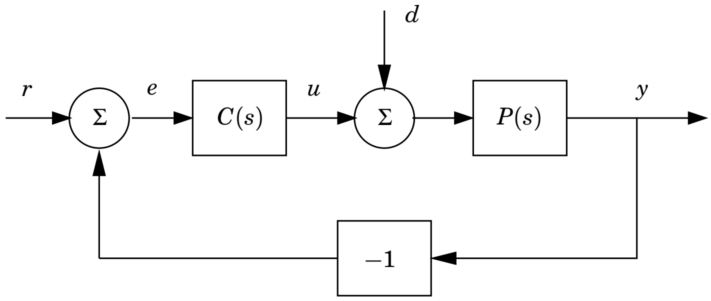

Figure 4.8 Block diagram of a first order system with a PI controller.

A linear model of a first order system can be described by the transfer function

$$P(s) = \frac{b}{s+a} \tag{4.13}$$

The system thus has two parameters. These parameters can be determined from physical consideration or from a step response test on the system. A step test will also reveal if it is reasonable to model a system by a first order model.

To have no steady state error a controller must have integral action. It is therefore natural to use a PI controller which has the transfer function

$$C(s) = k + \frac{k_i}{s} \tag{4.14}$$

A block diagram of the system is shown in Figure 4.8. The loop transfer function of the system is

$$L(s) = P(s)C(s) = \frac{kbs + k_i b}{s(s+a)} = \frac{n_L(s)}{d_L(s)} \tag{4.15}$$

The transfer function of the closed system from reference r to output y is given by

$$\frac{Y(s)}{R(s)} = \frac{P(s)C(s)}{1+P(s)C(s)} = \frac{n_L(s)}{d_L(s)+n_L(s)} = \frac{b(ks+k_i)}{s^2+(a+bk)s+bk_i}$$

The closed loop system is of second order and its characteristic polynomial is

$$d_L(s) + n_L(s) = s^2 + (a+bk)s + bk_i. \tag{4.16}$$

The poles of the closed loop system can be given arbitrary values by choosing the parameters k and k_i properly. Intuition about the effects of the parameters can be obtained from the mass-spring-damper analogy as was done in Section 4.2 and we find that integral gain k_i corresponds to stiffness and that proportional gain k corresponds to damping.

It is convenient to re-parameterize the problem so that the characteristic polynomial becomes

$$s^2 + 2\zeta\omega_0 s + \omega_0^2 \tag{4.17}$$

Identifying the coefficients of s in the polynomials (4.16) and (4.17) we find that the controller parameters are given by

$$\begin{aligned} k &= \frac{2\zeta\omega_0 - a}{b} \\ k_i &= \frac{\omega_0^2}{b} \end{aligned} \tag{4.18}$$

Since the design method is based on choosing the poles of the closed loop system it is called pole placement. Instead of choosing the controller parameters k and k_i we now select ζ and ω_0. These parameters have a good physical interpretation. The parameter ω_0 determines the speed of response and ζ determines the shape of the response. Controllers often have parameters that can be tuned manually. For a PI controller it is customary to use the parameters k and k_i. When a PI controller is used for a particular system, where the model is known, it is much more practical to use other parameters. If the model can be approximated by a first order model it is very convenient to have ω_0 and ζ as parameters. We call this performance related parameters because they are related directly to the properties of the closed loop system.

If the parameters ω_0 and ζ are known the controller parameters are given by (4.18). We will now discuss how to choose these parameters.

Behavior of Second Order Systems

We will first consider a second order system with the transfer function

$$G(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}. \tag{4.19}$$

This is a normalized transfer function of a second order system without zeros. The step responses of systems with different values of ζ are shown in Figure 4.9 The figure shows that parameter ω_0 essentially gives a time scaling. The response is faster if ω_0 is larger. The shape of the response

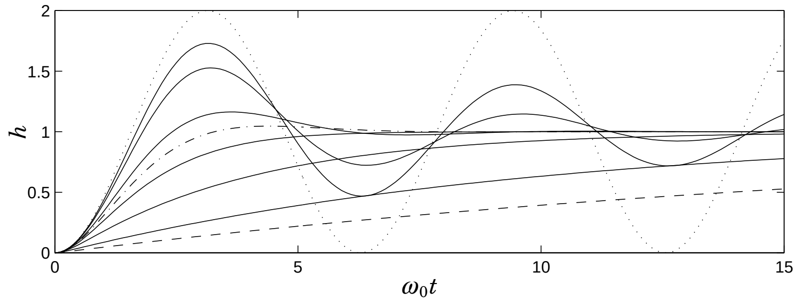

Figure 4.9 Step responses h for the system (4.19) with the transfer function $G(s) = \frac{\omega_0^2}{s^2+2\zeta\omega_0 s+\omega_0^2}$ for $\zeta = 0$ (dotted), 0.1, 0.2, 0.5, 0.707 (dash dotted), 1, 2, 5 and 10 (dashed).

is determined by ζ. The step responses have an overshoot of

$$M = \begin{cases} e^{-\dfrac{\pi\zeta}{\sqrt{1-\zeta^2}}} & \text{for } |\zeta| < 1 \\ 1 & \text{for} \zeta \geq 1 \end{cases}$$

For $\zeta < 1$ the maximum overshoot occurs at

$$t_{max} = \frac{2\pi}{\omega_0\sqrt{1-\zeta^2}}$$

There is always an overshoot if $\zeta < 1$. The maximum decreases and is shifted to the right when ζ increases and it becomes infinite for $\zeta = 1$ when the overshoot disappears. In most cases it is desirable to have a moderate overshoot which means that the parameter ζ should be in the range of 0.5 to 1. The value $\zeta = 1$ gives no overshoot.

Behavior of Second Order Systems with Zeros

We will now consider a system with the transfer function

$$G(s) = \frac{\omega_0}{\beta}\frac{s+\beta\omega_0}{s^2+2\zeta\omega_0 s+\omega_0^2} \tag{4.20}$$

Notice that the transfer function has been parameterized so that the steady state gain $G(0)$ is one. Step responses for this transfer function for different values of β are shown in Figure 4.10. The figure shows that

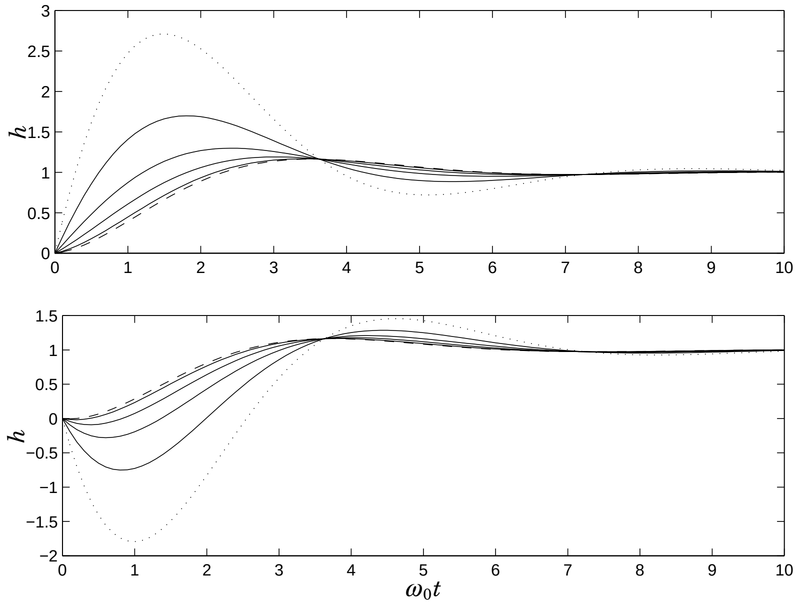

Figure 4.10 Step responses h for the system (4.19) with the transfer function $G(s) = \frac{\omega_0(s+\beta\omega_0)}{\beta(s^2+2\zeta\omega_0 s+\omega_0^2)}$ for $\omega_0 = 1$ and $\zeta = 0.5$. The values for $\beta = 0.25$ (dotted), 0.5 1, 2, 5 and 10 (dashed), are shown in the upper plot and $\beta = -0.25$, -0.5 -1, -2, -5 and -10 (dashed) in the lower plot.

the zero introduces overshoot for positive β and an undershoot for negative β. Notice that the effect of β is most pronounced if β is small. The effect of the zero is small if $|\beta| > 5$. Intuitively it it appears that systems with negative values of β, where the output goes in the wrong direction initially, are difficult to control. This is indeed the case as will be discussed later. Systems with this type of behavior are said to have inverse response. The behavior in the figures can be understood analytically. The transfer function $G(s)$ can be written as

$$G(s) = \frac{\omega_0}{\beta}\frac{s+\beta\omega_0}{s^2+2\zeta\omega_0 s+\omega_0^2} = \frac{\omega_0^2}{s^2+2\zeta\omega_0 s+\omega_0^2} + \frac{1}{\beta}\frac{s\omega_0}{s^2+2\zeta\omega_0 s+\omega_0^2}$$

Let $h_0(t)$ be the step response of the transfer function

$$G_0(s) = \frac{\omega_0^2}{s^2+2\zeta\omega_0 s+\omega_0^2}$$

It follows from (4.20) that the step response of $G(s)$ is

$$h(t) = h_0(t) + \frac{1}{\beta\omega_0}\frac{dh_0(t)}{dt} \tag{4.21}$$

It follows from this equation that all step responses for different values of β go through the point where dh_0/dt is zero. The overshoot will increase for positive β and decrease for negative β. It also follows that the effect of the zero is small if $|\beta|$ is large. The largest magnitude of dh/dt is approximately $0.4\omega_0/2.7$, which implies that the largest value of the second term is approximately $0.4/\beta$. The term is thus less than 8% if $|\beta|$ is larger than 5.

Notice in Figure 4.10 that the step response goes in the wrong direction initially when β is negative. This phenomena is called inverse response, can also be seen from (4.21). When β is negative the transfer function (4.20) has a zero in the right half plane. Such are difficult to control and they are called non-minimum phase system, see Section 3.5. Several physical systems have this property, for example level dynamics in steam generators (Example 3.18, hydro-electric power stations (Example 3.17), pitch dynamics of an aircraft (Example 3.19) and vehicles with rear wheel steering.

The Servo Problem

Having developed insight into the behavior of second order systems with zeros we will return to the problem of PI control of first order systems. We will discuss selection of controller parameters for the servo problem where the main concern is that the output should follow the reference signal well. The loop transfer function of the system is given by (4.15) and the transfer function from reference r to output y is

$$G_{yr} = \frac{Y(s)}{R(s)} = \frac{P(s)C(s)}{1+P(s)C(s)} = \frac{n_L(s)}{n_D(s)+n_L(s)} = \frac{(a+bk)s+bk_i}{s^2+(a+bk)s+bk_i}$$

Choosing control parameters to give the characteristic polynomial (4.17) we find as before that the controller parameters are given by (4.18) and the transfer function above becomes

$$\frac{Y(s)}{R(s)} = \frac{(a+bk)s+bk_i}{s^2+(a+bk)s+bk_i} = \frac{2\zeta\omega_0 s+\omega_0^2}{s^2+2\zeta\omega_0 s+\omega_0^2} \tag{4.22}$$

Comparing this transfer function with the transfer function (4.20) we find that

$$\beta = 2\zeta$$

This implies that parameter β is in the range of 1 to 2 for reasonable choices of ζ. Comparing with Figure 4.10 shows that the system has a significant overshoot. This can be avoided by a simple modification of the controller.

Avoiding the Overshoot - Systems with two degrees of freedom

The controller used in Figure 4.8 is based on error feedback. The control signal is related to the reference and the output in the following way

$$u(t) = k(r(t) - y(t)) + k_i \int_0^t (r(\tau) - y(\tau))d\tau \tag{4.23}$$

The reason for the overshoot is that the controller reacts quite violently on a step change in the reference. By changing the controller to

$$u(t) = -ky(t) + k_i \int_0^t (r(\tau) - y(\tau))d\tau \tag{4.24}$$

we obtain a controller that is reacting much less violent to changes in the reference signal. Taking Laplace transforms of this controller we get

$$U(s) = -kY(s) + \frac{k_i}{s}(R(s) - Y(s)) \tag{4.25}$$

Combining this equation with the equation (4.13) which describes the process we find that

$$\frac{Y(s)}{R(s)} = \frac{bk_i}{s^2 + (a + bk)s + bk_i} = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \tag{4.26}$$

and we obtain a transfer function from reference r to output y which does not have a zero, compare with (4.20).

The controller given by (4.23) is said to have error feedback because all control actions are based on the error $e = r - y$. The controller given by (4.24) is said to have two degrees of freedom (2DOF) because the signal path from reference r to control u is different from the signal path from output y to control u. Figure 4.11 shows block diagrams of the systems. The transfer function (4.26) is the standard transfer function for a second order system without zeros, its step responses are shown in Figure 4.9.

The Regulation Problem

It will now be investigated how the parameters ω_0 and ζ should be chosen for the regulation problem. In this problem the main concern is reduction

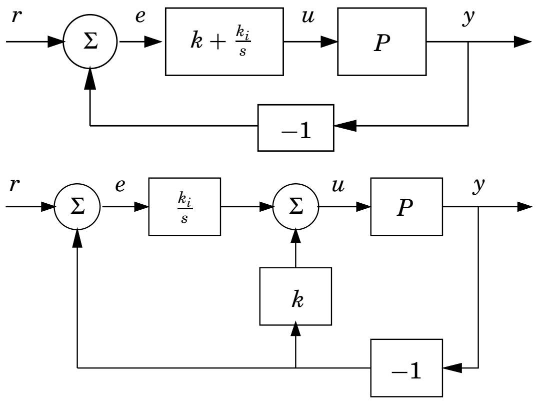

Figure 4.11 Block diagrams of a system with a conventional PI controller (above) and a PI controller having two degrees of freedom (below).

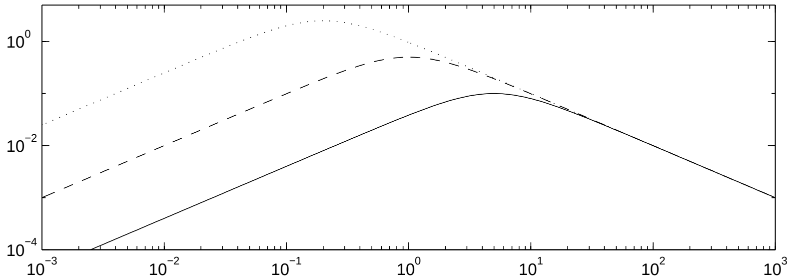

Figure 4.12 Gain curves for the transfer function from load disturbance to process output for $b = 1$, $\zeta = 1$ and $\omega_0 = 0.2$ dotted, $\omega_0 = 1.0$, dashed and $\omega_0 = 5$ full.

of load disturbances. Consider the system in Figure 4.8, the transfer function from load disturbance d to output y is

$$
\begin{aligned}
G_{yd}(s) = \frac{Y(s)}{D(s)} &= \frac{P(s)}{1 + P(s)C(s)} = \frac{s}{s^2 + (a + bk)s + bk_i} \\
&= \frac{bs}{s^2 + 2\zeta\omega_0 s + \omega_0^2} = \frac{b}{\omega_0}\frac{\omega_0 s}{s^2 + 2\zeta\omega_0 s + \omega_0^2}
\end{aligned}
$$

We will first consider the effect of parameter ω_0. Figure 4.12 shows the gain curves of the Bode diagram for different values of ω_0. The figure shows that disturbances of high and low frequencies are reduced significantly and that the disturbance reduction is smallest for frequencies around ω_0, they may actually be amplified. The figure also shows that the disturbance rejection at low frequencies is drastically influenced by the parameter ω_0 but that the reduction of high frequency disturbances is virtually independent of ω_0. It is easy to make analytical estimates because we have

$$G_{yd}(s) \approx \frac{bs}{\omega_0^2} = \frac{s}{bk_i}$$

for small s, where the second equality follows from (4.18). It follows from this equation that it is highly desirable to have a large value of ω_0. A large value of ω_0 means that the control signal has to change rapidly. The largest permissible value of ω_0 is typically determined by how quickly the control signal can be changed, dynamics that was neglected in the simple model (4.13) and possible saturations. The integrated error for a unit step disturbance in the load disturbance is

$$IE = \int_0^\infty e(t)dt = \lim_{s\to 0} E(s) = \lim_{s\to 0} G_{yd}\frac{1}{s} = \frac{b}{\omega_0^2} = \frac{1}{bk_i}$$

The largest value of $|G_{yd}(i\omega)|$ is

$$\max |G_{yd}(i\omega)| = |G_{yd}(i\omega_0)| = \frac{b}{2\zeta\omega_0}$$

The closed loop system obtained with PI control of a first order system is of second order. Before proceeding we will investigate the behavior of second order systems.

4.5 Control of Second Order Systems

We will now discuss control of systems whose dynamics can approximately be described by differential equations of second order. Such an approximation is reasonable if the storage of mass, momentum and energy can be captured by two state variables. Typical examples are

- Position of car on the road
- Motion control systems
- Stabilization of satellites
- Electric systems where energy is stored in two elements
- Levels in two connected tanks
- Pressure in two connected vessels

- Simple bicycle models

The general transfer function for a process of second order is

$$P(s) = \frac{b_1 s + b_2}{s^2 + a_1 s + a_2} \tag{4.27}$$

In some cases we will consider the special case when $b_1 = 0$.

PD control

We will first design a PD control of the process

$$P(s) = \frac{b}{s^2 + a_1 s + a_2}$$

A PD controller with error feedback has the transfer function

$$C(s) = k + k_d s$$

The loop transfer function is

$$L(s) = P(s)C(s) = \frac{bk_d s + bk}{s^2 + a_1 s + a_2} = \frac{n_L(s)}{d_L(s)}$$

The closed loop transfer function from reference to output is

$$\begin{aligned}\frac{Y(s)}{R(s)} &= \frac{PC}{1+PC} = \frac{n_L(s)}{n_D(s) + n_L(s)} = \frac{b(k_d s + k)}{s^2 + a_1 s + a_2 + b(k_d s + k)}\\ &= \frac{b(k_d s + k)}{s^2 + (a_1 + bk_d)s + a_2 + bk}\end{aligned}$$

The closed loop system is of second order and the controller has two parameters. The characteristic polynomial of the closed loop system is

$$s^2 + (a_1 + bk_d)s + a_2 + bk \tag{4.28}$$

Matching this with the standard polynomial

$$s^2 + 2\zeta\omega_0 s + \omega_0^2$$

we get

$$\begin{aligned}k &= \frac{\omega_0^2 - a_2}{b}\\ k_d &= \frac{2\zeta\omega_0 - a_1}{b}\end{aligned} \tag{4.29}$$

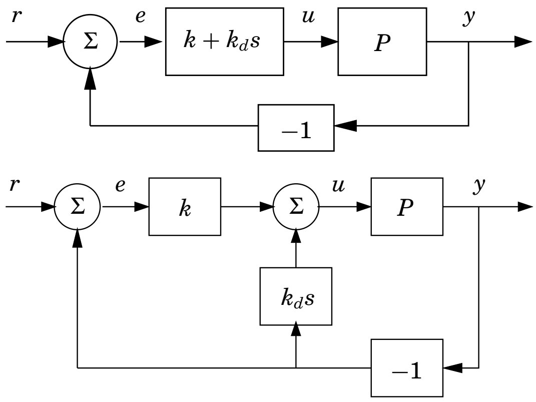

Figure 4.13 Block diagrams of system with PD control based on error feedback (above) and with a PD controller with two degrees of freedom (below). Compare with Figure 4.11.

The closed loop transfer function from reference to output becomes

$$\frac{Y(s)}{R(s)} = \frac{PC}{1+PC} = \frac{(2\zeta\omega_0 - a_1)s + \omega_0^2 - a_2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}$$

Notice that there will be a steady state error unless $a_2 = 0$. The steady state error is small if $|a_2| << \omega_0^2$. Also notice that the zero in the numerator may cause overshoot. To avoid this the controller based on error feedback can be replaced with the following controller

$$U(s) = k(R(s) - Y(s)) - k_d s Y(s) \tag{4.30}$$

which has two degrees of freedom. The transfer function from reference to output for the closed loop system then becomes

$$\frac{Y(s)}{R(s)} = \frac{\omega_0^2 - a_2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}$$

Notice that this transfer function does not have a zero. Block diagrams for the system with error feedback and with two degrees of freedom are shown in Figure 4.13.

PI Control

Next we will investigate what can be achieved with PI control of the process given by (4.27). Let the PI controller have the transfer function

$$C(s) = k + \frac{k_i}{s} = \frac{ks + k_i}{s}$$

The loop transfer function becomes

$$L(s) = P(s)C(s) = \frac{(ks + k_i)(b_1 s + b_2)}{s^3 + a_1 s^2 + a_2 s} = \frac{n_L(s)}{d_L(s)}$$

The characteristic polynomial is

$$n_L(s) + d_L(s) = s^3 + (a_1 + kb_1)s^2 + (a_2 + kb_2 + k_i b_1) + b_2 k_i$$

Identifying the coefficients of this equation with the desired characteristic polynomial

$$(s^2+2\zeta\omega_0 s+\omega_0^2)(s+\alpha\omega_0) = s^3+(\alpha+2\zeta)\omega_0 s^2+(1+2\alpha\zeta)\omega_0^2 s+\alpha\omega_0^3 \quad (4.31)$$

we obtain

$$\begin{aligned} a_1 + b_1 k &= (\alpha + 2\zeta)\omega_0 \\ a_2 + b_1 k_i + b_2 k &= (1 + 2\alpha\zeta)\omega_0^2 \\ b_2 k_i &= \alpha\omega_0^3 \end{aligned}$$

Since there are three equations and only two unknowns the problem cannot be solved in general. To have a solution we can let ω_0 be a free parameter. If $b_1 = 0$ and $b_2 \neq 0$ the equation then has the solution

$$\begin{aligned} \omega_0 &= \frac{a_1}{\alpha + 2\zeta} \\ k &= \frac{(1 + 2\alpha\zeta)\omega_0^2 - a_2}{b_2} \\ k_i &= \frac{\alpha\omega_0^3}{b_2} \end{aligned} \quad (4.32)$$

The parameter ω_0 which determines the response time is thus uniquely given by the process dynamics. When $b_1 \neq 0$ the parameter ω_0 is instead the real solution to the equation

$$\alpha b_1^2\omega_0^3 - (1 + 2\alpha\zeta)b_1 b_2\omega_0^2 + (\alpha + 2\zeta)b_2^2\omega_0 + a_2 b_1 b_2 - a_1 b_2^2 = 0.$$

and the controller parameters are given by

$$\begin{aligned} k &= \frac{(\alpha + 2\zeta)\omega_0 - a_1}{b_1} \\ k_i &= \frac{\alpha\omega_0^3}{b_2} \end{aligned}$$

In both cases we find that with PI control of a second order system there is only one choice of ω_0 that is possible. The performance of the closed loop system is thus severely restricted when a PI controller is used.

PID Control

Assume that the process is characterized by the second-order model

$$P(s) = \frac{b_1 s + b_2}{s^2 + a_1 s + a_2} \tag{4.33}$$

This model has four parameters. It has two poles that may be real or complex, and it has one zero. This model captures many processes, oscillatory systems, and systems with right half-plane zeros. The right half-plane zero can also be used as an approximation of a time delay. Let controller be

$$U(s) = k(bR(s) - Y(s)) + \frac{k_i}{s}(R(s) - Y(s)) + k_d s(cR(s) - Y(s))$$

The loop transfer function is

$$L(s) = \frac{(k_d s^2 + ks + k_i)(b_1 s + b_2)}{s(s^2 + a_1 s + a_2)} = \frac{n_L(s)}{d_L(s)}$$

The closed-loop system is of third order with the characteristic polynomial

$$\begin{aligned} d_L(s) + n_L(s) &= s(s^2 + a_1 s + a_2) + (b_1 s + b_2)(k_d s^2 + ks + k_i) \\ &= (1 + b_1 k)s^3 + (a_1 + b_1 k + b_2 k_d)s^2 + (a_2 + b_1 k_i + b_2 k)s + b_2 k_i \\ &= (1 + b_1 k)\Big(s^3 + \frac{a_1 + b_1 k + b_2 k_d}{1 + b_1 k}s^2 + \frac{a_2 + b_1 k_i + b_2 k}{1 + b_1 k}s + \frac{b_2 k_i}{1 + b_1 k}\Big) \end{aligned}$$

A suitable closed-loop characteristic equation of a third-order system is

$$(s + \alpha\omega_0)(s^2 + 2\zeta\omega_0 s + \omega_0^2)$$

Equating coefficients of equal power in s in this polynomial with the normalized characteristic polynomial gives

$$\begin{aligned} \frac{a_1 + b_2 k_d + b_1 k}{1 + b_1 k_d} &= (\alpha + 2\zeta)\omega_0 \\ \frac{a_2 + b_2 k + b_1 k_i}{1 + b_1 k_d} &= (1 + 2\alpha\zeta)\omega_0^2 \\ \frac{b_2 k_i}{1 + b_1 k_d} &= \alpha\omega_0^3 \end{aligned}$$

This is a set of linear equations in the controller parameters. The solution is straightforward but tedious and is given by

$$
\begin{aligned}
k &= \frac{a_2b_2^2 - a_2b_1b_2(\alpha+2\zeta)\omega_0 - (b_2 - a_1b_1)(b_2(1+2\alpha\zeta)\omega_0^2 + \alpha b_1\omega_0^3)}{b_2^3 - b_1b_2^2(\alpha+2\zeta)\omega_0 + b_1^2b_2(1+2\alpha\zeta)\omega_0^2 - \alpha b_1^3\omega_0^3}\\
k_i &= \frac{(-a_1b_1b_2 + a_2b_1^2 + b_2^2)\alpha\omega_0^3}{b_2^3 - b_1b_2^2(\alpha+2\zeta)\omega_0 + b_1^2b_2(1+2\alpha\zeta)\omega_0^2 - \alpha b_1^3\omega_0^3}\\
k_d &= \frac{-a_1b_2^2 + a_2b_1b_2 + b_2^2(\alpha+2\zeta)\omega_0 - b_1b_2\omega_0^2(1+2\alpha\zeta) + b_1^2\alpha\omega_0^3}{b_2^3 - b_1b_2^2(\alpha+2\zeta)\omega_0 + b_1^2b_2(1+2\alpha\zeta)\omega_0^2 - \alpha b_1^3\omega_0^3}
\end{aligned} \tag{4.34}
$$

The transfer function from set point to process output is

$$
G_{yr}(s) = \frac{(b_1s + b_2)(ck_ds^2 + bks + k_i)}{(s+\alpha\omega_0)(s^2 + 2\zeta\omega_0s + \omega_0^2)}
$$

The parameters b and c have a strong influence on shape of the transient response of this transfer function.

The transfer function from load disturbance to process output is

$$
G_{yd} = \frac{b_1s^2 + b_2s}{(s+\alpha\omega_0)(s^2 + 2\zeta\omega_0s + \omega_0^2)}
$$

These formulas are useful because many processes can be approximately described by the transfer function (4.27). We illustrate this with an example.

EXAMPLE 4.1—OSCILLATORY SYSTEM WITH RHP ZERO
Consider a system with the transfer function

$$
P(s) = \frac{1-s}{s^2+1}
$$

This system has one right half-plane zero and two undamped complex poles. The process is difficult to control.

$$
s^3 + 2s^2 + 2s + 1.
$$

(4.34) gives a PID controller with the parameters $k = 0$, $k_i = 1/3$, and $k_d = 2/3$. Notice that the proportional gain is zero. □

We will give an example that illustrates that there are situations where a PID controller can be much better than a PI controller.

EXAMPLE 4.2—PID CAN BE MUCH BETTER THAN PI

Consider a process described by

$$P(s) = \frac{k_v}{s(1+sT)} e^{-sT_d} \tag{4.35}$$

where the time delay T_d is much smaller than the time constant T. Since the time constant T is small it can be neglected and the design can be based on the second order model

$$P(s) \approx \frac{k_v}{s(1+sT)} \tag{4.36}$$

A PI controller for this system can be obtained from Equation (4.32) and we find that a closed loop system with the characteristic polynomial (4.31) can be obtained by choosing the parameter ω_0 equal to $1/(\alpha + 2\zeta)T$. Since $T_d << T$ it follows that $\omega_0 T_d << 1$ and it is reasonable to neglect the time delay.

If the approximation (4.36) it is possible to find a PID controller that gives the closed loop characteristic polynomial with arbitrarily large values of ω_0. Since the real system is described by (4.35) the parameter ω_0 must be chosen so that the approximation (4.36) is valid. This requires that the product $\omega_0 T_d$ is not too large. It can be demonstrated that the approximation is reasonable if $\omega_0 T_d$ is smaller than 0.2.

Summarizing we find that it is possible to obtain the characteristic polynomial (4.31) with both PI and PID control. With PI control the parameter ω_0 must be chosen as $1/(\alpha + 2\zeta)T$. With PID control the parameter instead can be chosen so that the product $\omega_0 T_d < 1$ is small, e.g. 0.2 or less. With PI control the response speed is thus determined by T and with PID control it is determined by T_d. The differences can be very significant. Assume for example that $T = 100$, $T_d = 1$, $\alpha = 1$ and $\zeta = 0.5$. Then we find that with $\omega_0 = 0.005$ with PI control and $\omega_0 = 0.1$ with PID control. This corresponds to a factor of 200 in response time. This will also be reflected in a much better disturbance attenuation with PID control. □

4.6 Control of Systems of High Order*

The method for control design used in the previous sections can be characterized in the following way. Choose a controller of given complexity, PD, PI or PID and determine the controller parameters so that the closed loop characteristic polynomial is equal to a specified polynomial. This technique is called pole placement because the design is focused on achieving a closed loop system with specified poles. The zeros of the transfer function from reference to output can to some extent be influenced by choosing a controller with two degrees of freedom. We also observed that the complexity of the controller reflected the complexity of the process. A PI controller is sufficient for a first order system but a PID controller was required for a second order system. Choosing a controller of too low order imposed restrictions on the achievable closed loop poles. In this section we will generalize the results to systems of arbitrary order. This section also requires more mathematical preparation than the rest of the book.

Consider a system given by the block diagram in Figure 4.8. Let the process have the transfer function

$$P(s) = \frac{Y(s)}{U(s)} = \frac{b(s)}{a(s)} = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \ldots + b_n}{s^n + a_1 s^{n-1} + \ldots + a_n} \tag{4.37}$$

where $a(s)$ and $b(s)$ are polynomials. A general controller can be described by

$$f(s)U(s) = -g(s)Y(s) + h(s)R(s) \tag{4.38}$$

where $f(s)$, $g(s)$ and $h(s)$ are polynomials. The controller given by (4.38) is a general controller with two degrees of freedom. The transfer function from measurement signal y to control signal u is $-g(s)/f(s)$ and the transfer function from reference signal r to control signal u is $h(s)/f(s)$. For a system with error feedback we have $g(s) = h(s)$. Elimination of $U(s)$ between Equations (4.37) and (4.38) gives

$$\big(a(s)f(s) + b(s)g(s)\big)Y(s) = b(s)h(s)R(s) + b(s)f(s)D(s) \tag{4.39}$$

The closed loop has the characteristic polynomial

$$c(s) = a(s)f(s) + b(s)g(s) \tag{4.40}$$

Notice that this only depends on the polynomials $f(s)$ and $g(s)$. The design problem can be stated as follows: Given the polynomials $a(s)$, $b(s)$ and $c(s)$ find the polynomials $f(s)$ and $g(s)$ which satisfies (4.40). This is a well known mathematical problem. It will be shown in the next section that the equation always has a solution if the polynomials $a(s)$ and $b(s)$ do not have any common factors. If one solution exists there are also infinitely many solutions. This is useful because it makes it possible to introduce additional constraints. We may for example require that the controller should have integral action.

A Naive Solution

To obtain the solution to the design problem the equation (4.40) must be solved. A simple direct way of doing this is to introduce polynomials f and g with arbitrary coefficients, writing equating coefficients of equal powers of s, and solving the equations. This procedure is illustrated by an example.

EXAMPLE 4.3—GENERAL POLE PLACEMENT
Consider a process with the transfer function

$$P(s) = \frac{1}{(s+1)^2}$$

Find a controller that gives a closed loop system with the characteristic polynomial

$$(s^2 + as + a^2)(s + a)$$

(4.40) becomes

$$(s+1)^2 f + g = (s^2 + as + a^2)(s + a) = s^3 + 2as^2 + 2a^2 s + a^3$$

One solution is

$$\begin{aligned} f &= 1 \\ g &= s^3 + (2a-1)s^2 + (2a^2-2)s + a^3 - 1 \end{aligned}$$

but there are also other solutions e.g.

$$\begin{aligned} f &= s + 2a - 2 \\ g &= (2a^2 - 4a + 3)s + a^3 - 2a + 2 \end{aligned}$$

□

The Diophantine Equation

The naive solution of (4.40) hides many interesting aspects of the problem. The equation (4.40) is a classical equation which has been studied extensively in mathematics. To discuss this equation we will use more mathematics than in most parts of the book. We will also change to a more formal style of presentation. This is a nice illustration of the fact that control is a field where many branches of mathematics are useful.

We will start by observing that polynomials belong to a mathematical object called a ring. This means that they can be multiplied and added, and that there are units: the zero polynomial for addition and the polynomial 1 for multiplication. Division of polynomials does not always give a polynomial, but quotient and remainders are defined. Integers are other objects that also is a ring. To develop some insight we will first explore two examples.

EXAMPLE 4.4—AN EQUATION IN INTEGERS
Consider the following equation

$$3x + 2y = 1,$$

where x and y are integers. By inspection we find that $x = 1$ and $y = -1$ is a solution. We also find that if we have a solution other solutions can be obtained by adding 2 to x and subtracting 3 from y. The equation thus has infinitely many solutions. □

EXAMPLE 4.5—AN EQUATION IN INTEGERS
Consider the equation

$$6x + 4y = 1,$$

where x and y are integers. This equation cannot have a solution because the left hand side is an even number and the right hand side is an odd number. □

EXAMPLE 4.6—AN EQUATION IN INTEGERS
Consider the equation

$$6x + 4y = 2,$$

where x and y are integers. Dividing the right hand side by 2 we obtain the equation in Example 4.4 □

These examples tell most about the (4.40) when a, b, f, g and c belong to a ring. To be precise we have the following result.

THEOREM 4.1—EUCLID'S ALGORITHM
Let a, b, and c be polynomials with real coefficients. Then the equation

$$ax + by = c \tag{4.41}$$

has a solution if and only if the greatest common factor of a and b divides c. If the equation has a solution x_0 and y_0 then $x = x_0 - bn$ and $y = y_0 + an$, where n is an arbitrary integer, is also a solution.

PROOF 4.1
We will first determine the largest common divisor of the polynomials a and b by a recursive procedure. Assume that the degree of a is greater than or equal to the degree of b. Let $a^0 = a$ and $b^0 = b$. Iterate the equations

$$\begin{aligned} a^{n+1} &= b^n \\ b^{n+1} &= a^n \text{ mod b}^\text{n} \end{aligned}$$

until $b^{n+1} = 0$. The greatest common divisor is then b^n. If a and b are co-prime we have $b^n = 1$. Backtracking we find that

$$ax + by = b^n$$

where the polynomials x and y can be found by keeping track of the quotients and the remainders in the iterations. When a and b are co-prime we have

$$ax + by = 1$$

and the result is obtained by multiplying x and y by c. When a and b have a common factor it must be required that the largest common divisor of a and b is also a factor of c. Dividing the equation with this divisor we are back to the case when a and b are co-prime. □

Since the proof has only used addition, multiplication, quotients and remainders it follows that the results holds for any ring.

An Algorithm

The following is a convenient way of organizing the recursive computations. With this method we also obtain the minimum degree solution to the homogeneous equation.

$$\begin{aligned} ax + by &= 1 \\ au + bv &= 0 \end{aligned} \tag{4.42}$$

where g is the greatest common divisor of a and b and u and v are the minimal degree solutions to the homogeneous equation These equations can be written as

$$\begin{pmatrix} x & y \\ u & v \end{pmatrix} \begin{pmatrix} a & 1 & 0 \\ b & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x & y \\ 0 & u & v \end{pmatrix}$$

The solution to Equation (4.42) can be obtained by transforming the matrix

$$A^0 = \begin{pmatrix} a & 1 & 0 \\ b & 0 & 1 \end{pmatrix}$$

by row operations to a matrix where the 21 element is zero. This can be done recursively as follows. Assume that $\deg a$ is greater than or equal to $\deg b$, exchange the rows if this is not the case. Form the following recursion.

$$A^{n+1} = \begin{pmatrix} A_{21}^n & A_{22}^n & A_{23}^n \\ r^n & A_{12}^n - q^n A_{22}^n & A_{13}^n - q^n A_{23}^n \end{pmatrix}$$

where $q^n = A_{11}^n$ div $\mathrm{A}_{21}^{\mathrm{n}}$ and $r_n = A_{11}^n$ div $\mathrm{A}_{21}^{\mathrm{n}}$. Proceed until $A_{21}^{n+1} = 0$. It follows from Euclid's algorithm that A_{11}^n is the greatest common divisor of a and b and that a and b are co-prime if $A_{11}^n = 1$. The equation (4.41) then has a solution if A_{11}^n is a factor of c.

System Theoretic Consequences

The following result is an immediate consequence of Euclid's algorithm, Theorem 4.1.

THEOREM 4.2—CONTROLLER PARAMETERIZATION
Consider a system with a rational transfer function $P = b/a$. Let $C_0 = g_0/f_0$ be a controller which gives a closed loop system with the characteristic polynomial c. Then all controllers which give a closed loop system with the characteristic polynomial c are given by

$$C = \frac{g_0 + qa}{f_0 - qb}$$

where q is an arbitrary polynomial.

PROOF 4.2
The loop transfer function obtained with the controller C is

$$L = PC = \frac{b(g_0 + qa)}{a(f_0 - qb)}$$

we have

$$1 + L = \frac{a(f_0 - qb) + b(g_0 + qa)}{a(f_0 - qb)} = \frac{af_0 + bg_0}{a(f_0 - qb)} = \frac{c}{a(f_0 - qb)}$$

which shows that the characteristic polynomial is c. Let $C = g/f$ be any controller that gives the characteristic polynomial c it follows that

$$af + bg = c$$

and it follows from Theorem 4.1 that $f = f_0 - bq$ and $g = g_0 + aq$. □

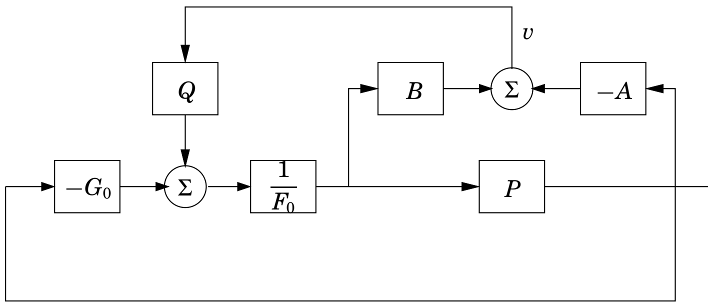

Figure 4.14 Block diagram that illustrates the Youla-Kučera parameterization theorem. If $C = G_0/F_0$ stabilizes the system $P = B/A$, then the controller shown in the block diagram also stabilizes the system for all stable rational functions Q.

This theorem is useful because it characterizes all controllers that give specified closed loop poles. Since the theorem tells that there are many solutions we may ask if there are some solutions that are particularly useful. It is natural to look for simple solutions. It follows from Theorem 4.2 that there is one controller where $\deg f < \deg b$, i.e. a controller of lowest order, and another where $\deg g < \deg a$, a controller with highest pole excess.

Youla-Kučera Parameterization

Theorem 4.2 characterizes all controllers that give a closed loop system with a given characteristic polynomial. We will now derive a related result that characterizes all stabilizing controllers. To start with we will introduce another representation of a transfer function.

DEFINITION 4.1—STABLE RATIONAL FUNCTIONS
Let $a(s)$ be a polynomial with all zeros in the left half plane and $b(s)$ an arbitrary polynomial. The rational function $b(s)/a(s)$ is called a stable rational function.

□

Stable rational functions are also a ring. This means that Theorem 4.1 also holds for rational functions. A fractional representation of a transfer function P is

$$P = \frac{B}{A}$$

where A and B are stable rational transfer functions. We have the following result.

THEOREM 4.3—YOULA-KUČERA REPRESENTATION
Consider a process with the transfer function $P = B/A$, where A and B are stable rational functions that are co-prime, let $C_0 = G_0/F_0$ be a fractional representation of a controller that stabilizes P, all stabilizing controllers are then given by

$$C = \frac{G_0 + QA}{F_0 - QB} \tag{4.43}$$

where Q is an arbitrary stable rational transfer function.

PROOF 4.3
The loop transfer function obtained with the controller C is

$$L = PC = \frac{B(G_0 + QA)}{A(F_0 - QB)}$$

we have

$$1 + L = \frac{A(F_0 - QB) + B(G_0 + QA)}{A(F_0 - QB)} = \frac{AF_0 + BG_0}{A(F_0 - QB)}$$

Since the rational function $AF_0 + BG_0$ has all its zeros in the left half plane the closed loop system is stable. Let $C = G/F$ be any controller that stabilizes the closed loop system it follows that

$$AF + BG = C$$

is a stable rational function with all its zeros in the left half plane. Hence

$$\frac{A}{C}F + \frac{B}{C}G = 1$$

and it follows from Theorem 4.1 that

$$F = F_0 - \frac{B}{C}Q = F_0 - B\bar{Q}$$
$$G = G_0 - \frac{A}{C}Q = G_0 - A\bar{Q}$$

where Q is a stable rational function because C has all its zeros in the left half plane. □

It follows from Equation (4.43) that the control law can be written as

$$\frac{U}{Y} = -\frac{G}{F} = -\frac{G_0 + QA}{F_0 - QB}$$

or

$$F_0 U = -G_0 Y + Q(BU - AY)$$

The Youla-Kučera parameterization theorem can then be illustrated by the block diagram in Figure 4.14. Notice that the signal v is zero. It therefore seems intuitively reasonable that a feedback based on this signal cannot make the system unstable.

4.7 Summary

In this section we started by investigating some simple control systems. A systematic method for analysis and design was developed. The closed loop system was first represented by a block diagram. The behavior of each block was represented by a transfer function. The relations between the Laplace transforms of all signals could be derived by simple algebraic manipulations of the transfer functions of the blocks. An interesting feature of using Laplace transforms is that systems and signals are represented in the same way. The analysis gave good insight into the behavior of simple control systems and how its properties were influenced by the poles and zeros of the closed loop system. The results can also be developed using differential equations but it is much simpler to use Laplace transforms and transfer functions. This is also the standard language of the field of control.

To design a controller we selected a controller with given structure, PI or PID. The parameters of the controller were then chosen to obtain a closed loop system with specified poles, or equivalently specified roots of the characteristic equation. This design method was called pole placement. The design methods were worked out in detail for first and second order systems but we also briefly discussed the general case. To find suitable closed loop poles we found that it was convenient to introduce standard parameters to describe the closed loop poles. Results that guide the intuition of choosing the closed loop poles were also developed.

The analysis was based on simplified models of the dynamics of the process. The example on cruise control in Section 4.2 indicated that it was not necessary to know some parameters accurately. One of the amazing properties of control systems is that they can often be designed based on simple models. This will be justified in the next chapter.