Control Theory Summary

1. Preliminary knowledge

Ever since mankind existed, they have tried to control things. During the past decennia, people have actually gotten quite good at it. So now it's our turn to take a look at how we can control systems.

1.1 Introduction to systems

To examine systems, we first need to know a few things: What are systems? And what kinds of systems are there? That's what we will look at first.

1.1.1 What is a system?

What is a system? A **system** is a combination of components that (together) perform a certain objective. Systems usually vary in time. This happens during a so-called **process**. A process is a set of controlled actions leading towards a certain end.

Another way to look at a system, is by considering it as some 'magic box'. You simply insert a certain input, denoted as $u(t)$ and out comes a certain output, denoted as $y(t)$. So a system merely takes an input and gives an output, according to a specified way.

1.1.2 System variables

Systems usually have variables. There are two important kinds of variables.

- First, there are the **manipulated variables**. They are controlled (set) by the controller. They are thus the input of the system. (Think of the deflection you give to the steering wheel of your car.)

- Second, there are the **controlled variables**. These variables are influenced by the manipulated variables. They are usually (but not always) the output of the system. (Think of the heading/position of the car.)

1.1.3 System types

There are many ways in which to distinguish systems. Let's have a look at a few important ones.

- Open/closed loop systems: In **closed-loop systems** (also called **feedback control systems**), the output is compared with the input. The result of this comparison is then used to change the output. In other words, in these systems the output effects the control action. Although such systems are usually relatively complicated, they handle disturbances rather well. However, in **open-loop system** the output does not effect the control action. These systems are usually quite simple. They aren't very good at dealing with disturbances though.

- Scalar/multivariable systems: In **scalar** systems there is only one input $u(t)$ and one output $y(t)$. **Multivariable systems** have multiple inputs $u_i(t)$ or multiple outputs $y_i(t)$ (or both).

- Linear/nonlinear systems: Let's suppose $y_1(t)$ is the respone of a system to input $u_1(t)$. Similarly, $y_2(t)$ is the respone to $u_2(t)$. Now examine the input signal $u(t) = c_1 u_1(t) + c_2 u_2(t)$. If the output of the system then is $y(t) = c_1 y_1(t) + c_2 y_2(t)$ (for all combinations of c_1, c_2, $u_1(t)$ and $u_2(t)$), then the system is **linear**. Otherwise, it is **nonlinear**.

1.2 The Laplace transform

A very important tool we need to evaluate systems, is the Laplace transform. This transform has been discussed in the differential equations course already. Although we will quickly go over it, it is advised that you have some preliminary knowledge on the Laplace transform already.

1.2.1 Definition of the Laplace transform

Let's suppose we have a function $f(t)$. The Laplace transform of $f(t)$, written as $\mathscr{L}(f(t))$, is defined as

$$\mathscr{L}(f(t)) = \int_0^\infty f(t)e^{-st}dt. \tag{1.2.1}$$

For example, the Laplace transform of the function $f(t) = at$ (with a constant) is

$$\mathscr{L}(at) = \int_0^\infty ate^{-st}dt = \left[-\frac{at}{s}e^{-st}\right]_0^\infty + \int_0^\infty \frac{a}{s}e^{-st}dt = 0 + \left[-\frac{a}{s^2}e^{-st}\right]_0^\infty = \frac{a}{s^2}. \tag{1.2.2}$$

Here s is a complex variable, which we will take a closer look at later. We denote the **inverse Laplace transform** as $\mathscr{L}^{-1}(F(s))$. For example, we have $\mathscr{L}^{-1}(a/s^2) = at$.

We can find Laplace transforms of a lot of functions with the above mentioned integral. This is, however, rather boring. And it has been done a lot of times already. Therefore so-called Laplace tables exist. These tables contain lists of function $f(t)$ and their corresponding Laplace transform $\mathscr{L}(f(t))$. If you want to take the Laplace transform of a function, I suggest you use such a table. (One can be found in the book on Control Theory, on pages 17 and 18. You can also find one on the page of the differential equations course.)

1.2.2 Special functions

There are several special functions, which have a rather basic Laplace transform. Let's have a look at them. First, there is the **ramp function** $r(t)$, defined as

$$r(t) = \begin{cases} 0 & \text{if } t < 0, \\ t & \text{if } t \geq 0. \end{cases} \tag{1.2.3}$$

The Laplace transform of this function is $\mathscr{L}(r(t)) = 1/s^2$. Second, there is the **unit step function** $u(t)$. This function is, in fact, the derivative of the ramp function. It is defined as

$$u(t) = \begin{cases} 0 & \text{if } t < 0, \\ 1 & \text{if } t \geq 0. \end{cases} \tag{1.2.4}$$

Its Laplace transform is given by $\mathscr{L}(u(t)) = 1/s$. And third, there is the **unit impulse function** $\delta(t)$. This, in turn, is the derivative of the unit step function. It is defined such that

$$\delta(t) = \begin{cases} 0 & \text{if } t \neq 0, \\ \infty & \text{if } t = 0, \end{cases} \qquad \text{but also} \qquad \int_{-\infty}^\infty \delta(t)\, dt = 1. \tag{1.2.5}$$

Its Laplace transform is given by $\mathscr{L}(\delta(t)) = 1$.

1.2.3 Special rules for derivatives

When using the Laplace transform, we can also apply several rules. Some of these rules are actually quite handy. For example, the rules for integrals and derivatives. Let's suppose we have a function $f(x)$ and know its Laplace transform $F(s) = \mathscr{L}(f(t))$. What are the Laplace transforms of the derivate and integral of $f(x)$? Well, they are given by

$$\mathscr{L}\left(\frac{df}{dt}\right) = sF(s) - f(0) \qquad \text{and} \qquad \mathscr{L}\left(\int f(t)\, dt\right) = \frac{F(s) + f(0)}{s}. \qquad (1.2.6)$$

In this course, we usually assume that $f(0) = 0$. In that case the relations above simplify to

$$\mathscr{L}\left(\frac{df}{dt}\right) = sF(s) \qquad \text{and} \qquad \mathscr{L}\left(\int f(t)\, dt\right) = \frac{F(s)}{s}. \qquad (1.2.7)$$

So we see that, if we take a derivative of $f(t)$, we should multiply its Laplace transform by s. Similarly, if we integrate $f(t)$, we should divide its Laplace transform by s. This is an interesting rule, which you should remember. By the way, if you look at the previous paragraph (about the special functions), you see that this rule holds.

1.2.4 Change-of-variable rules

But there are more special rules for the Laplace transform. We can, for example, apply a change-of-variable. Let's once more suppose that we know $F(s) = \mathscr{L}(f(t))$. Then what is the Laplace transform of $f(at + b)$? Well, it is

$$\mathscr{L}(f(at + b)) = \frac{e^{\frac{b}{a}}}{a} F\left(\frac{s}{a}\right). \qquad (1.2.8)$$

We can also apply a change-of-variable the other way around. Again we have $F(s) = \mathscr{L}(f(t))$. But what is the inverse Laplace transform of $F(as + b)$? This then is

$$\mathscr{L}^{-1}(F(as + b)) = \frac{e^{-\frac{b}{a}t}}{a} f\left(\frac{t}{a}\right). \qquad (1.2.9)$$

By the way, it is also important to know that the Laplace transform is a **linear operator**. This means that

$$\mathscr{L}(c_1 f_1(t) + c_2 f_2(t)) = c_1 \mathscr{L}(f_1(t)) + c_2 \mathscr{L}(f_2(t)). \qquad (1.2.10)$$

1.2.5 Roots and poles of a function

Let's suppose we have a function $F(s)$. For example, let's examine

$$F(s) = \frac{K(s + 2)(s + 10)}{s(s + 1)(s + 5)^2(s + 10)(s^2 + 9)}. \qquad (1.2.11)$$

We know that s is a complex variable. So we can examine this function on all points in the complex plane. The points where this function is analytic (where a unique derivative exists) are called **ordinary points**. Other points are called **singular points**. Singular points can be classified as either **zeroes**, **poles** or neither.

To find the zeroes, we need to look at the numerator. This should be zero. For $F(s)$, the value $s = -2$ is a zero. To find the poles, we need to look at the denominator. Now this should be zero. So, for $F(s)$, the poles are $s = 0$, $s = -1$, $s = -5$, $s = 3i$ and $s = -3i$. One may initially think that $s = -10$ is both a pole and a zero. However, it is neither. $s = -10$ is just a singular point.

Poles and zeroes also have an **order**. This is the amount of times which they occur. In our example, the pole $s = -1$ occurred just once. (There was only one factor $(s + 1)$.) Its order is thus 1. The order of $s = -5$, however, is two. This is because the factor $(s + 5)$ occurred twice.

We have actually missed a zero of $F(s)$. Let's ask ourselves, what happens if $s \to \infty$? In this case the constant coefficients within the brackets don't really matter anymore. (For example, as $s \to \infty$, then $s + 5 \approx s$.) So $F(s) \to Ks^2/s^7 = K/s^5$. Thus $s = \infty$ is a zero of fifth order.

It is important to note that zeroes and poles aren't always real numbers. They can also be complex numbers, as we have already seen. An interesting fact, however, is that complex zeroes/poles always come in pairs. If $s = a + bi$ is a pole, then its complex conjugate $\bar{s} = a - bi$ is also a pole. The same goes for zeroes.

1.2.6 Finding the inverse Laplace transform using partial-fraction expansion

Let's suppose that we have a function $F(s)$, and want to find the inverse Laplace transform $\mathscr{L}^{-1}(F(s))$. We can get a Laplace transform table, and hope that $F(s)$ is in there. But if it is not, what do we do then? Panic? Rather not. Instead, we can apply a partial-fraction expansion to $F(s)$. For that, we first have to rewrite $F(s)$ in the form

$$F(s) = \frac{B(s)}{A(s)} = K \frac{(s - z_1)^{r_{z1}}(s - z_2)^{r_{z2}} \ldots (s - z_m)^{r_{zm}}}{(s - p_1)^{r_{p1}}(s - p_2)^{r_{p2}} \ldots (s - p_n)^{r_{pn}}}. \tag{1.2.12}$$

Here $z_1, z_2, \ldots, z_m$ are the zeroes of $F(s)$, having orders $r_{z1}, r_{z2}, \ldots, r_{zn}$, respectively. The poles p_i are defined similarly. Now we want to rewrite $F(s)$. How can we do that? Well, to simplify matters, we first examine the case where all poles have order 1 (so $r_{p1} = r_{p2} = \ldots = r_{pn} = 1$). We can then write $F(s)$ as

$$F(s) = \frac{a_1}{(s - p_1)^{r_{p1}}} + \frac{a_2}{(s - p_2)^{r_{p2}}} + \ldots + \frac{a_n}{(s - p_n)^{r_{pn}}}. \tag{1.2.13}$$

The coefficients $a_1, a_2, \ldots, a_n$ are called the **residues** at the poles. The question now is, how can we find these residues? Well, to find a residue a_k, we simply multiply both sides by the factor $(s - p_k)$. We then insert $s = p_k$ into the equation. Due to this, a lot of terms will vanish. We remain with

$$a_k = \left((s - p_k) \frac{B(s)}{A(s)} \right)_{s=p_k}. \tag{1.2.14}$$

It may seem like the right side of the equation will be zero. This is, however, not the case, because $A(s)$ has a factor $(s - p_k)$ as well. These factors will thus cancel each other out. So then we can simply find the residue a_k.

But what do we do if there is a pole p_k with order greater than 1? Let's, for example, say that this order is $r_{pk} = 3$. In this case we have to include more terms in $F(s)$. 3 more terms, to be exact. Let's call the part of $F(s)$ representing this pole $F_k(s)$. We then have

$$F_k(s) = \frac{a_{k_1}}{(s - p_k)} + \frac{a_{k_2}}{(s - p_k)^2} + \frac{a_{k_3}}{(s - p_k)^3}. \tag{1.2.15}$$

How do we find the residues a_{k_1}, a_{k_2} and a_{k_3} now? We can find a_{k_3} in a rather familiar way. We multiply both sides of the equation by $(s - p_k)^3$. We then get

$$(s - p_k)^3 F_k(s) = a_{k_1}(s - p_k)^2 + a_{k_2}(s - p_k) + a_{k_3}. \tag{1.2.16}$$

By inserting $s = p_k$ we can find a_{k_3}. But how do we find the other coefficients? To find them, we differentiate with respect to s. Differentiating once gives

$$\frac{d}{ds}\left((s - p_k)^3 F_k(s) \right) = 2a_{k_1}(s - p_k) + a_{k_2}. \tag{1.2.17}$$

By inserting $s = p_k$ we can find a_{k_2}. To find a_{k_1}, we have to differentiate once more. (And if the order of the pole is even higher, you have to differentiate even more times to find the other coefficients.)

In the end you should have the partial-fraction expansion of $F(s)$. And then it is quite easy to find the inverse Laplace transform $f(t)$.

2. Mathematically modelling systems

It is time to take an actual look at systems now. But to do that, we need to have some way to display them. There are several ways to do this. We'll examine these methods in this chapter.

2.1 Block diagrams

2.1.1 Block diagram notations

As we now know, systems manipulate variables. They take input variables $u(t)$ and do all sort of stuff with them, to find the output variables $y(t)$. They can, for example, add them up ($y(t) = u_1(t) + u_2(t)$). They can subtract them ($y(t) = u_1(t) - u_2(t)$). They can multiply them ($y(t) = Ku(t)$). They can even integrate them ($y(t) = \int_0^t u(\tau)\, d\tau + y(0)$).

But how can we write/display this? One common-used method to do that, is by using **block diagrams**. Examples of parts of such diagrams are shown in figure 2.1. As you can see, addition and subtraction of two signals is denoted by a circle with a plus or a minus next to it. Other operations (like multiplying and integrating) are usually noted by a square, with the corresponding factor in it.
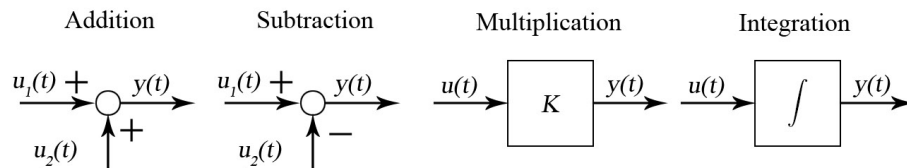

Figure 2.1: Examples of block diagram parts.

There is, however, a slight problem. Evaluating the integral in a block diagram is often rather troubling. Luckily, there is an easy solution to this problem. We simply examine block diagrams in the Laplace domain. Adding up and subtracting remains unchanged, as well as multiplying by constants. However, we can now simply replace the integral by a factor $1/s$. (Remember that integrating meant dividing by s in the Laplace domain.) That simplifies things quite a lot.

2.1.2 A basic closed-loop system

Let's examine a basic **closed-loop system**, as is displayed in figure 2.2. As you can see, it has a so-called **feedback loop**. The system also has an **input signal** $U(s)$ and an **output signal** $Y(s)$. (Note that, since we are now working in the Laplace domain, we use capital letters.)
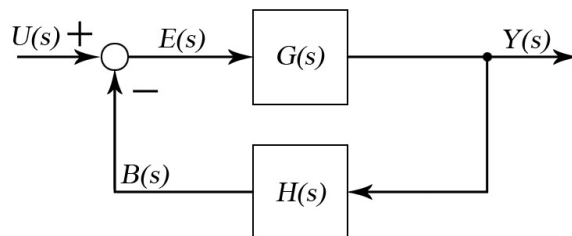

Figure 2.2: A basic closed loop system.

Let's take a look at what's happening in this system. The system recieves a certain input signal $U(s)$. It compares this with the so-called **feedback signal** $B(s)$. (In fact, it takes the difference.) This gives the **actuating error signal** $E(s)$. This error signal is then passed on to some **control function** $G(s)$, which returns the output $Y(s)$. However, the function $Y(s)$ isn't only outputted. It is also transformed by the function $H(s)$ and fed back to the system as the feedback signal.

2.1.3 Controller types

In the closed-loop system, there was the control function $G(s)$. Given an error $E(s)$ (or equivalently, $e(t)$), this function determines the output $Y(s)$ (or similarly, $y(t)$). The manner in which it does this, is called the **control action**. We, as designers of the system, are supposed to define this control action. There are several controllers/control actions that are very common. Let's take a look at some of them.

- In the **proportional controller** (the **P-controller**), the output is proportional to the error. This means that $y(t) = K_p e(t)$. In the Laplace domain this becomes $G(s) = Y(s)/E(s) = K_p$. The variable K_p is called the **proportional gain**.

- In the **proportional + integral controller** (the **PI-controller**) we have

$$u(t) = K_p e(t) + \frac{K_p}{T_i} \int_0^t e(t)\,dt, \qquad \text{or, in the Laplace domain,} \qquad G(s) = \frac{U(s)}{E(s)} = K_p \left(1 + \frac{1}{T_i s}\right).$$
$$(2.1.1)$$

  The variable T_i is called the **integral time**.

- In the **proportional + derivative controller** (the **PD controller**) we have

$$u(t) = K_p e(t) + K_p T_d \frac{de(t)}{dt}, \qquad \text{or, in the Laplace domain,} \qquad G(s) = \frac{U(s)}{E(s)} = K_p \left(1 + T_d s\right).$$
$$(2.1.2)$$

  Surprisingly, T_d is called the **derivative time**.

- Finally, in the **proportional + integral + derivative controller** (the **PID controller**) we combine the previous two cases. We then have

$$u(t) = K_p e(t) + \frac{K_p}{T_i} \int_0^t e(t)\,dt + K_p T_d \frac{de(t)}{dt}. \tag{2.1.3}$$

  In the Laplace domain this of course becomes

$$G(s) = \frac{U(s)}{E(s)} = K_p \left(1 + \frac{1}{T_i s} + T_d s\right). \tag{2.1.4}$$

2.2 Transfer functions

Performing calculations with block diagrams is rather difficult. That's why we now examine another modelling method: the transfer function.

2.2.1 Definition of the transfer function

The **transfer function** $F(s)$ of a system is defined as the ratio between the output and the input in the Laplace domain. So,

$$F(s) = \frac{Y(s)}{U(s)}. \tag{2.2.1}$$

By the way, to find the above transfer function, we do have to assume that all initial conditions are zero. Let's try to find the transfer function for the system of figure 2.2. First, we can find that

$$Y(s) = E(s)G(s) = (U(s) - B(s))G(s) = (U(s) - H(s)Y(s))G(s). \tag{2.2.2}$$

Solving for $Y(s)/U(s)$ gives the rather important **closed-loop transfer function**

$$F(s) = \frac{G(s)}{1 + G(s)H(s)}. \tag{2.2.3}$$

By the way, from the transfer function, we can also find the **order** of the system. The order of the system is the highest power of s in the denominator of $F(s)$. So if, for example, $F(s) = (a_0 + a_1 s)/(b_0 + b_1 s + b_2 s^2)$, then the order is 2. (This is due to the $b_2 s^2$ term.)

2.2.2 Transfer functions of parts

We can also find transfer functions for parts of systems. We once more examine the closed-loop system. There is, for example, the **open-loop transfer function**. This is the transfer function if the loop would not be closed. It is thus

$$\text{open loop transfer function} = \frac{B(s)}{E(s)} = G(s)H(s). \tag{2.2.4}$$

Similarly, the **feedforward transfer function** is defined as the multiplication of all blocks directly between the input and the output. It is thus

$$\text{feedforward transfer function} = \frac{Y(s)}{E(s)} = G(s). \tag{2.2.5}$$

2.2.3 The impulse response function

Another very important function is the **impulse response function** $f(t)$. This is the inverse Laplace transform of the transfer function $F(s)$. (Note that $f(t)$ is not in the Laplace domain, but is an actual function.) To find it, we simply have to find $\mathscr{L}^{-1}(F(s))$.

Why is the impulse response function so important? To find that out, we suppose that the input of the system is only a unit impulse function (thus $u(t) = \delta(t)$). We then have $U(s) = 1$ and thus also $F(s) = Y(s)$. In other words $f(t)$ will then be the output of the system! The impulse response function is thus the output of the system when the input is a unit impulse.

And why is this so important? Well, suppose that we know this input response function. It is then possible (using some mathematical tricks) to find the response (output) of the system for any input function. And that is of course what we want. We will, however, investigate response analysis in later chapters.

2.2.4 Response of a system to initial conditions

Let's suppose we have a system with a known transfer function $F(s)$. Also the input function $u(t)$ (or equivalently, $U(s)$) is known. We can then determine the output, using the method described above. (Although we will go into greater detail in later chapters.) However, we have made one rather important assumption: the initial conditions are zero. But what should we do if this is not the case? We then write our transfer function $F(s)$ as

$$F(s) = \frac{Y(s)}{U(s)} = \frac{b_0 + b_1 s + b_2 s^2 + \dots}{a_0 + a_1 s + a_2 s^2 + \dots}. \tag{2.2.6}$$

It can be shown that the above equation corresponds to the differential equation

$$b_0 u(t) + b_1 \dot{u}(t) + b_2 \ddot{u}(t) + \ldots = a_0 y(t) + a_1 \dot{y}(t) + a_2 \ddot{y}(t) + \ldots. \tag{2.2.7}$$

We now assume that the input $u(t)$ is zero (thus also $U(s) = 0$), but that the initial conditions are not zero. This simplifies the above differential equation. If we then again transform it to the Laplace domain, we can find that

$$Y(s) = \frac{a_1 y(0) + a_2 (s y(0) + \dot{y}(0)) + a_3 (s^2 y(0) + s \dot{y}(0) + \ddot{y}(0)) + \ldots}{a_0 + a_1 s + a_2 s^2 + \ldots}. \tag{2.2.8}$$

Find the inverse Laplace transform of this function and you will know the response $y(t)$ to the initial conditions. By the way, if there are both initial conditions and a non-zero input, you can use the principle of superposition. Simply add both solutions up.

2.2.5   Transforming a block diagram to a transfer function

Let's suppose we have a system with some very complicated block diagram. How do we find the transfer function? This is, in fact, a rather difficult problem. However, there are a few rules you can use to simplify things.

- When two blocks $G_1(s)$ and $G_2(s)$ are placed in sequence (with no branching points or joining points in between), you can replace them by one block. This block is the multiple $G_1(s)G_2(s)$ of the two blocks.

- When two blocks $G_1(s)$ and $G_2(s)$ are placed parallel to each other (again without branching/joining), and are then joined together, you can replace them by one block. This block then contains their sum $G_1(s) + G_2(s)$ or their difference $G_1(s) - G_2(s)$. (This depends on whether there is a plus or a minus sign at the joining of the blocks.)

- When there is a closed-loop part in the system (as shown in figure 2.2), you can replace it by one block. This block will then contain $G(s)/(1 + G(s)H(s))$.

It often occurs that you can't simplify your block diagram any further with the above rules. This is because branching/joining points are interfering. When this is the case, you have to move these branching/joining points. However, this often causes the values of blocks to change. You may wonder, what do these blocks change to? Well, there are two rules which you should always keep in mind.

- The product of all blocks around any loop must stay the same.

- The product of all blocks in the feedforward direction (the path from input to output) must stay the same.

By applying these rules, you should be able to simplify the entire block diagram to one block. This block then contains the transfer function of the entire system. And this is exactly what you wanted to find.

2.3   The state-space representation

A more modern method is the state-space method. Especially computers are rather good at using this method. It is also slightly less abstract than a transfer function, because it directly uses differential equations.

2.3.1 State variables and the state vector

Let's examine a system. This system has r inputs $u_1(t), \ldots, u_r(t)$ and m outputs $y_1(t), \ldots, y_m(t)$. For this system, we can define a set of n **state variables**. These state variables should completely describe the system. However, there may not be any redundancy: if we remove any of these variables, the system isn't completely defined anymore. If this is the case, then these n variables form the **state** of a system.

Working with n state variables can be a bit troublesome. We therefore usually put them in a vector

$$\mathbf{x}(t) = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}^T. \tag{2.3.1}$$

This vector is the so-called **state vector**. Similarly, we can define an **input vector** $\mathbf{u}(t)$ and an **output vector** $\mathbf{y}(t)$.

You may wonder, how many state variables do we need? This amount n (also called the **order** of the system) is equal to the number of integrators in our system. In fact, the n state variables $x_1, \ldots, x_n$ are the outputs of these integrators. This may sound rather vague at the moment. However, it will be clearer once we deal with a practical example. We will do that soon.

2.3.2 State-space equations

So, we have these n variables. They aren't just random variables. They are governed by rules. From these rules, we should be able to derive n differential equations, being

$$\dot{x}_i(t) = f_i(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r, t), \tag{2.3.2}$$

for every i from 1 to n. We should also be able to derive the outputs $y_i(t)$ from these state variables. So,

$$y_i(t) = g_i(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r, t), \tag{2.3.3}$$

for every i from 1 to m. This all seems rather difficult. However, things often simplify a lot. Most of the times the functions f_i and g_i are linear. (We then have a so-called **linear system**.) In this case, we can write

$$\dot{\mathbf{x}}(t) = A(t)\mathbf{x}(t) + B(t)\mathbf{u}(t) \qquad \text{and} \qquad \mathbf{y}(t) = C(t)\mathbf{x}(t) + D(t)\mathbf{u}(t). \tag{2.3.4}$$

And things get even better now. For the problems we will consider, the matrices A, B, C and D are constant in time as well. (The system is then **time-invariant**.) So we have

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t) \qquad \text{and} \qquad \mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t). \tag{2.3.5}$$

The first of these equations is the **state equation**. The second is the **output equation**. The matrices A, B, C and D have names too. A is the $(n \times n)$ **state matrix**, B is the $(n \times r)$ **input matrix**, C is the $(m \times n)$ **output matrix** and D is the $(m \times r)$ **direct transmission matrix**.

2.3.3 Transforming a state-space system to a transfer function

There is a link between state-space systems and transfer functions. We can transform a state-space system to a transfer function. To do that, we first have to take the Laplace transform of equation (2.3.5). This gives us

$$s\mathbf{X}(s) - \mathbf{x}(0) = A\mathbf{X}(s) + B\mathbf{U}(s) \qquad \text{and} \qquad Y(s) = C\mathbf{X}(s) + D\mathbf{U}(s). \tag{2.3.6}$$

From the first equation, we can solve for $\mathbf{X}(s)$. We then insert it into the second one. This gives us

$$\mathbf{Y}(s) = (C(sI - A)^{-1}B + D)\mathbf{U}(s). \tag{2.3.7}$$

Since $\mathbf{Y}(s) = F(s)\mathbf{U}(s)$, we find that

$$F(s) = C(sI - A)^{-1}B + D. \tag{2.3.8}$$

Just like D, also $F(s)$ is an $m \times r$ matrix. You may wonder how that is possible. We wanted a transfer function, and now we have a transfer function matrix. That is because we now have multiple inputs and outputs. In fact, the matrix entry $F_{ij}(s)$ gives the transfer function for the output i, due to input j. (This, of course, follows directly from the vector equation $\mathbf{Y}(s) = F(s)\mathbf{U}(s)$.) If our system only has one input and one output, then $F(s)$ will just be a normal transfer function.

You also may have wondered, does the inverse of $(sI - A)$ always exist? In fact, it doesn't. If $\det(sI - A) = 0$, then this inverse does not exist. $F(s)$ is then not defined. From the linear algebra course, we remember that this is only the case when s is an eigenvalue of A. However, from the part on transfer functions we also remember an important fact: $F(s)$ is not defined when s is a pole. What can we conclude from this? It means that the eigenvalues of A are the poles of $F(s)$!

2.3.4   A practical example of a state-space system

It is time to consider a practical example. Let's consider the system in figure 2.3. It has one input $u(t)$, being the deflection of the rightmost point. As output, we want to have $\dot{y}_1(t)$ and $y_2(t)$. We want to model this system using the state-space method. To find the state-space system, we must find the matrices A, B, C and D. So how do we do that?
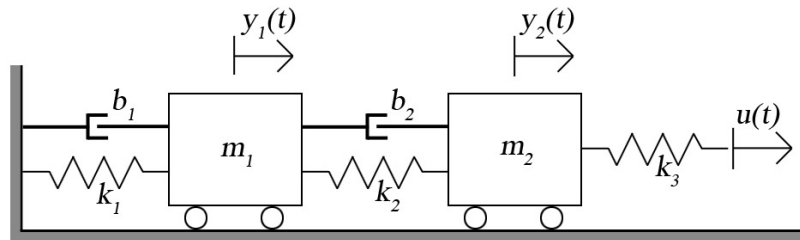

Figure 2.3: An example system.

The first step is by finding the equation of motion. For the first car we have

$$m_1 \ddot{y}_1 = -k_1 y_1 - k_2(y_1 - y_2) - b_1 \dot{y}_1 - b_2(\dot{y}_1 - \dot{y}_2). \tag{2.3.9}$$

Sometimes the signs in the above equation get confusing. If this troubles you, you can apply a trick. For example, let's look at the spring with stiffness k_1. Ask yourself, if $y_1 > 0$, will the force due to the spring be to the left or to the right? In this case the force will point to the left. (The spring is stretched out. It therefore pulls at the block.) However, y_1 is positive if it goes to the right. Since the directions are opposite, there must be a minus sign.

Similarly, we can find for the second car that

$$m_2 \ddot{y}_2 = -k_2(y_2 - y_1) - b_2(\dot{y}_2 - \dot{y}_1) - k_3(y_2 - u). \tag{2.3.10}$$

Now we want to put these equations into a state-space system. But what state variables should we define? And how many? Well, we can recall from earlier in this chapter that the amount of state variables was equal to the amount of integrators. How many integrators do we have in this case? We notice that there is a second derivative of $y_1(t)$. So to get $y_1(t)$ itself, we need two integrators. Similarly, we need two integrators for $y_2(t)$. We thus have four integrators. We should therefore define four state variables: x_1, x_2, x_3 and x_4.

But what state variables shall we define? A logical choice would be to define $x_1 = y_1$ and $x_2 = y_2$. But what about the other two variables? Well, you may have noticed that there are second derivatives in our equations. And in the state-space equations there are no second derivatives. How do we deal with this? We simply define $x_3 = \dot{y}_1$ and $x_4 = \dot{y}_2$. This implies that $\dot{x}_3 = \ddot{y}_1$ and $\dot{x}_4 = \ddot{y}_2$. Using these definitions, we can rewrite equations (2.3.9) and (2.3.10) to

$$\dot{x}_3 = \frac{-k_1 - k_2}{m_1} x_1 + \frac{k_2}{m_1} x_2 + \frac{-b_1 - b_2}{m_1} x_3 + \frac{b_2}{m_1} x_4, \tag{2.3.11}$$

$$\dot{x}_4 = \frac{k_2 x_1}{m_2} + \frac{-k_2 - k_3}{m_2} x_2 + \frac{b_2}{m_2} x_3 + \frac{-b_2}{m_2} x_4 + \frac{k_3}{m_2} u. \tag{2.3.12}$$

Now we are ready to find the state equation. It will be

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{-k_1 - k_2}{m_1} & \frac{k_2}{m_1} & \frac{-b_1 - b_2}{m_1} & \frac{b_2}{m_1} \\ \frac{k_2}{m_2} & \frac{-k_2 - k_3}{m_2} & \frac{b_2}{m_2} & \frac{-b_2}{m_2} \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{k_3}{m_2} \end{bmatrix} \mathbf{u}. \tag{2.3.13}$$

So we have found A and B! That's great news. Now only C and D remain. To find them, we have to find expressions for the outputs of the system, being $\dot{y}_1(t)$ and $y_2(t)$. This is, in fact, easier than it seems. We simply have $\dot{y}_1 = x_3$ and $y_2 = x_2$. This implies that

$$\mathbf{y} = C\mathbf{x} + D\mathbf{u} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \mathbf{u}. \tag{2.3.14}$$

And now we have found C and D as well. So this is how you find the state-space equations of a system. Just write down the equations of motion, define your state variables, and find the matrices A, B, C and D.

3. Analyzing systems subject to basic inputs

We now know how we can display systems. The next step is to analyze the systems. Let's subject systems to basic input functions. What can we then say about their behaviour?

3.1 Analyzing basic systems

This is where we start our journey into the wondrous world of stability. We begin by examining the definitions of stability, after which we examine some simple systems.

3.1.1 Definition of stability

Let's suppose we have a system. This system has zero input (so $u(t) = 0$). It may occur that the position/state of the system doesn't change in time. If this is the case, then the system is in a so-called **equilibrium position**.

Now let's give the system certain initial conditions. (We put it in a position other than the equilibrium position.) What will happen? This, in fact, depends on the system. Some systems converge back to an equilibrium position. These systems are called **stable**. Other systems oscillate about an equilibrium position, but don't converge to it. Such systems are **critically stable**. If the system diverges, then it is **unstable**.

3.1.2 First order systems

Let's examine a first order system. The transfer function of such a system can be written as

$$F(s) = \frac{1}{Ts + 1}. \tag{3.1.1}$$

We want to investigate the stability of this system. There is, however, a slight problem. Every system behaves differently to different input functions. So we can't give any general rules about the behaviour of a system.

However, we want to get at least some indication of the system behaviour. So we examine its response to some basic functions: the unit impulse function, the unit step function and the unit ramp function. Another basic function is the sine function. However, we will not examine the response of a system to periodic inputs now. We leave that for the chapter on frequency response analysis.

Let's suppose the input of the system is a unit impulse function. In this case $U(s) = 1$ and thus $Y(s) = F(s)$. We then find that

$$y(t) = \mathscr{L}^{-1}\left(F(s)\right) = \mathscr{L}^{-1}\left(\frac{1}{Ts + 1}\right) = \frac{1}{T}e^{-t/T}. \tag{3.1.2}$$

But what if the input is a unit step function? We then have $U(s) = 1/s$ and thus $Y(s) = F(s)/s$. The output of the system then becomes

$$y(t) = \mathscr{L}^{-1}\left(F(s)/s\right) = \mathscr{L}^{-1}\left(\frac{1}{Ts^2 + s}\right). \tag{3.1.3}$$

To find the inverse Laplace transform, we can apply a partial-fraction expansion. This would give us

$$y(t) = \mathscr{L}^{-1}\left(\frac{1}{Ts^2 + s}\right) = \mathscr{L}^{-1}\left(\frac{1}{s} - \frac{T}{Ts + 1}\right) = 1 - e^{-t/T}. \tag{3.1.4}$$

Finally, we examine the unit ramp function. We now have $U(s) = 1/s^2$ and thus $Y(s) = F(s)/s^2$. The output of the system now is

$$y(t) = \mathcal{L}^{-1}\left(F(s)/s^2\right) = \mathcal{L}^{-1}\left(\frac{1}{Ts^3 + s^2}\right) = \mathcal{L}^{-1}\left(\frac{1}{s^2} - \frac{T}{s} + \frac{T^2}{Ts+1}\right) = t - T + Te^{-t/T}. \quad (3.1.5)$$

3.1.3 The steady-state error

There are a couple of important parameters describing the behaviour of a system. The **error** $e(t)$ is the difference between the output and the input. So, $e(t) = u(t) - y(t)$. The **steady-state error** error e_{ss} is now defined as

$$e_{ss} = \lim_{t\to\infty} e(t). \quad (3.1.6)$$

This limit may seem rather annoying, because we first have to find $e(t)$. But luckily, we don't always have to! This is because we can use **initial/final value theorem**. This theorem states that

$$\lim_{t\to 0} f(t) = \lim_{s\to\infty} sF(s) - f(0) \qquad \text{and} \qquad \lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s) - f(0). \quad (3.1.7)$$

We can use the second part of this theorem. Our initial conditions are zero, so $e(0) = 0$. We also have $E(s) = U(s) - Y(s)$. This gives

$$e_{ss} = \lim_{t\to\infty} s(U(s) - Y(s)). \quad (3.1.8)$$

Now let's try to find the steady state error for the first order system of the previous paragraph. If the input is a unit impulse function or a unit step function, the steady-state error is zero. However, for the unit ramp function, this is not the case. This steady-state error is equal to

$$e_{ss} = \lim_{t\to\infty} u(t) - y(t) = \lim_{t\to\infty} t - (t - T - e^{-t/T}) = \lim_{t\to\infty} T + e^{-t/T} = T. \quad (3.1.9)$$

(Note that we don't have to apply the theorem we just showed. This is because we already have an expression for $e(t)$.) You may wonder, why does the system have an error when it is subject to a unit ramp function? We will investigate the reasons behind this later.

3.1.4 Second order systems

Now let's examine second order systems. Let's suppose we have a transfer function of the form

$$F(s) = \frac{c}{as^2 + bs + c}. \quad (3.1.10)$$

We can rewrite this to the so-called **standard form** of the second order system, being

$$F(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \quad (3.1.11)$$

where $\omega_n^2 = c/a$ and $2\zeta\omega_n = b/a$. By the way, ω_n is the **undamped natural frequency** and ζ is the **damping ratio**. If you have attended the vibrations course, both terms will be very familiar to you.

Let's suppose that the input is a unit step function. We thus have $Y(s) = F(s)/s$. What will the output $y(t)$ of the system be? If $0 < \zeta < 1$, then we are dealing with the **underdamped case**. In this case the solution will have oscillation. In fact, we can find that

$$y(t) = 1 - e^{-\zeta\omega_n t}\left(\cos\omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\omega_d t\right). \quad (3.1.12)$$

14

By the way, the variable $\omega_d = \omega_n \sqrt{1 - \zeta^2}$ is the **damped natural frequency**. If $\zeta = 1$, then we are dealing with the **critically damped case**. This time no oscillations occur. Instead, we have

$$y(t) = 1 - e^{-\omega_n t}(1 + \omega_n t). \tag{3.1.13}$$

Finally, there is the case $\zeta > 1$. We now have an **overdamped system**. The solution for $y(t)$ is

$$y(t) = 1 - \frac{1}{2\sqrt{\zeta^2 - 1}} \left(\frac{e^{-(\zeta - \sqrt{\zeta^2 - 1})\omega_n t}}{\zeta - \sqrt{\zeta^2 - 1}} - \frac{e^{-(\zeta + \sqrt{\zeta^2 - 1})\omega_n t}}{\zeta + \sqrt{\zeta^2 - 1}} \right). \tag{3.1.14}$$

This time no oscillations occur either. Instead, the system slowly converges to the equilibrium position.

3.1.5 Other important parameters

Next to the steady-state error, there are several other important parameters describing the response of a system. Their definition is most simple if the input of a system is a unit step function. (Thus $U(s) = 1/s$.) So we examine that case.

- The **delay time** t_d is the time required for the response to reach half of the final value. For a unit step function, we thus have $y(t_d) = 1/2$.

- The **rise time** t_r is the time required for the response to rise from either 10% to 90%, from 5% to 95% or from 0% to 100%. (Which one to use depends on the circumstances.)

- The **peak time** t_p is the time required for the response to reach the first peak. So, t_p is the smallest positive value for which $dy(t_p)/dt = 0$.

- The **maximum percent overshoot** is the percentage of overshoot at the first peak. So this is

$$\frac{y(t_p) - y(\infty)}{y(\infty)} \times 100\%. \tag{3.1.15}$$

- The **settling time** t_s is the time required for the response to reach and stay within a certain percentage of the final value. This percentage is usually 2% or 5%. We thus have, for the 2% case,

$$\frac{|y(t) - y(\infty)|}{|y(\infty)|} \leq 0.02 \qquad \text{for} \qquad t \geq t_s. \tag{3.1.16}$$

3.2 General rules for stability

We have previously examined first order and second order systems. Now let's examine systems in general. What can we then say about the stability of a system?

3.2.1 Finding the general solution

Let's examine a system with transfer function $F(s)$. We can write this function as

$$F(s) = K \frac{(s - z_1)(s - z_2) \dots (s - z_m)}{(s - p_1)(s - p_2) \dots (s - p_n)}, \tag{3.2.1}$$

where $z_1, \dots, z_m$ are the zeroes and $p_1, \dots, p_n$ are the poles. Now what will happen if this system is subject to a unit step function as input? To find that out, we will try to find the solution for $y(t)$. We know we have $Y(s) = F(s)/s$. We can use a partial-fraction expansion to rewrite this to

$$Y(s) = \frac{F(s)}{s} = \frac{a}{s} + \sum_{i=1}^{n} \frac{a_i}{s - p_i}. \tag{3.2.2}$$

Let's take a closer look at these poles p_i. They are either real or complex. However, if they are complex, then they always come in pairs of complex conjugates. So we can assume that there are q real poles $p_1, \ldots, p_q$ and r complex pole pairs, being $(\lambda_1 \pm \eta_1 i), \ldots, (\lambda_r \pm \eta_r i)$. (Note that $q + 2r = n$.) We can then find that

$$Y(s) = \frac{F(s)}{s} = \frac{a}{s} + \sum_{j=1}^{q} \frac{a_j}{s - p_j} + \sum_{k=1}^{r} \frac{b_k(s - \lambda_k) + c_k \eta_k}{s^2 - 2\lambda_k s + (\lambda_k^2 + \eta_k^2)}. \tag{3.2.3}$$

Finding the inverse Laplace transform gives us

$$y(t) = a + \sum_{j=1}^{q} a_j e^{p_j t} + \sum_{k=1}^{r} b_k e^{\lambda_k t} \cos \eta_k t + \sum_{k=1}^{r} c_k e^{\lambda_k t} \sin \eta_k t. \tag{3.2.4}$$

And this is the solution to our problem.

3.2.2 The importance of poles

It's very nice that we found a solution in the previous paragraph. But what can we conclude from it? Well, something very important! We can find whether $y(t)$ converges.

Let's suppose there is some pole of $Y(s)$ with a real part bigger than zero. (So either $p_j > 0$ for a real pole or $\lambda_k > 0$ for a complex pole.) In this case $y(t) \to \infty$ as $t \to \infty$. In other words, the system is unstable! Therefore, to ensure stability, the real part of all poles must be smaller than zero.

However, we can conclude more from this. Let's examine two poles p_i and p_j, one with a very negative real part λ_i, and one with an (also) negative real part λ_j closer to zero. Which one is more important? Well, the term $e^{\lambda_i t}$ (for the very negative pole) goes to zero very quickly as t increases. However, the term $e^{\lambda_j t}$ (for the less negative pole) stays relatively big. Therefore the pole p_j is more important. And in general we can say: the higher the real part of the pole, the more influential this pole is.

You may wonder, what is the effect of the complex part of the poles? Well, for that we examine the part $\sin \eta_k t$. We see that η_k is the angular frequency of this part. In other words, complex poles with a high value for η_k cause high frequency oscillations.

And finally, you may be wondering, don't zeroes have any effect? Well, you are right that zeroes aren't very important. However, they do have some effect. It can be shown that if a zero z_i is close to some pole p_j, then the residue a_j of that pole is small. This reduces the effect of the pole. What can we conclude from this? A zero near a pole reduces the effect of the pole.

3.2.3 System types

Now we know whether $y(t)$ converges. But if it converges, does it converge to the actual input? Or will there be a steady-state error? That's what we will examine shortly.

But first, we examine a system. In fact, we examine a basic feedback system, with unity feedback. (This means that $H(s) = 1$.) The transfer function of the system thus becomes

$$F(s) = \frac{Y(s)}{U(s)} = \frac{G(s)}{1 + G(s)}. \tag{3.2.5}$$

Let's take a closer look at this control function $G(s)$. In fact, let's write it as

$$G(s) = K \frac{(1 - s/z_1)(1 - s/z_2) \ldots (1 - s/z_m)}{s^N (1 - s/p_1)(1 - s/p_2) \ldots (1 - s/p_{n-N})}. \tag{3.2.6}$$

Here the values p_i are all nonzero. In other words, we define N as the number of times which the pole $s = 0$ occurs for $G(s)$. We now say that our entire feedback system is a **type N system**. So if $N = 0$

(and $s = 0$ is thus not a pole of $G(s)$), then we have a type 0 system. If $N = 1$, we have a type 1 system. And this continues for $N = 2, 3, \ldots$.

Now what's the use of this? Well, it will turn out that the system type will say something about the steady-state error of the system. And that's what we'll look at in the upcoming paragraph.

3.2.4 Effecting the steady-state error of the system

We are going to investigate the steady-state error of a basic feedback system. To do that, we must first find the error function $E(s)$. The transfer function is given by equation (3.2.5). The error function $E(s)$ will then be

$$E(s) = \frac{Y(s)}{G(s)} = \frac{U(s)}{1 + G(s)}. \tag{3.2.7}$$

To find the steady-state error, we can apply the final value theorem. It states that

$$e_{ss} = \lim_{s \to 0} \frac{sU(s)}{1 + G(s)}. \tag{3.2.8}$$

From equation (3.2.6) we can see that

$$\lim_{s \to 0} G(s) = \lim_{s \to 0} K/s^N. \tag{3.2.9}$$

Now let's suppose that the input function is $U(s) = 1/s^q$, for some value q. We then find that

$$e_{ss} = \lim_{s \to 0} \frac{1}{s^{q-1} + Ks^{q-1-N}}. \tag{3.2.10}$$

Great! Now what does it mean?

- If $q = 0$ (and we thus have a unit impulse function as input), then $s^{q-1} \to \infty$. This means $e_{ss} = 0$, regardless of the system.

- If $q = 1$ (meaning the input is a unit step function), then $s^{q-1} \to 1$. We can't say anything about the steady-state yet. So we have to examine Ks^{q-1-N}. If $N = q-1 = 0$, then $Ks^{q-1-N} = K$. This means that there will be a steady-state error. In fact, we find that $e_{ss} = 1/(1 + K)$. If, however, $N > 0$, we find that $Ks^{q-1-N} \to \infty$. This means that $e_{ss} = 0$. So there is no steady-state error.

- If $q \geq 2$, then $s^{q-1} \to 0$. So we again have to examine Ks^{q-1-N}. If $N > q - 1$, we find that $Ks^{q-1-N} \to \infty$. This means that $e_{ss} = 0$. So there is no steady-state error. If, however, $N > q-1$, then $e_{ss} \to \infty$. The system thus can't follow the input. Finally, if $N = q - 1$, then the steady-state error is $e_{ss} = 1/K$.

Let's summarize the results. If $N < q - 1$, there is an infinite steady-state error. If $N = q - 1$, there is a finite steady-state error. And if $N > q - 1$, there is no steady-state error. So to get rid of our steady-state error, we must increase N.

By the way, there is a very common mistake here. To find the type of the system, you should not examine the transfer function $F(s)$ of the system. Instead, investigate the feedforward transfer function $G(s)$. Only then will you find the right conclusions.

4. Root-locus plots

It's time we start designing systems ourselves. But how can we make sure our system gets the properties we want? In this chapter we'll investigate one method for that: the root-locus method.

4.1 General drawing rules

First we will examine the basic rules of the root-locus method. In the next part we will examine how they work.

4.1.1 The basic ideas

Suppose we have a (rather familiar) closed-loop system, as is shown in figure 4.1. We know that the transfer function is $F(s) = G(s)/(1 + G(s)H(s))$. We also know that the properties of the system strongly depend on the poles of $F(s)$. We would like to have some influence on these poles.
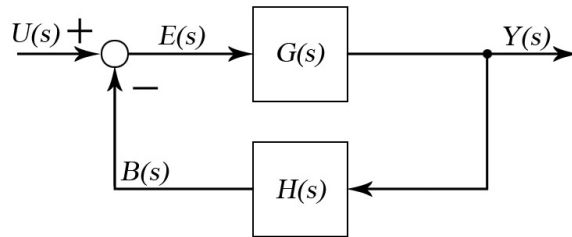

Figure 4.1: Again, a basic closed loop system.

Well, if we want to find the poles of $F(s)$, we should set its denominator to zero. So that gives us the equation

$$G(s)H(s) = K\frac{(s-z_1)(s-z_2)\ldots(s-z_m)}{(s-p_1)(s-p_2)\ldots(s-p_n)} = -1. \tag{4.1.1}$$

This equation is called the **characteristic equation**. Every s satisfying this equation is a pole of $F(s)$. By the way, the variables $z_1, \ldots, z_m$ and $p_1, \ldots, p_n$ in the above equation are the zeroes and poles, respectively, of $G(s)H(s)$. (Not to be confused with the zeroes and poles of $F(s)$.)

We assume that all these zeroes and poles of $G(s)H(s)$ are already defined. There is, however, one variable left for us to define, being K. Luckily it turns out that the poles of $F(s)$ (which we want to define) strongly depend on K. We therefore would like to do plot these poles as a function of K. In other words, we want to plot all s satisfying the above equation.

This is, however, more difficult than it seems. Especially because the above equation is a complex equation. So how do we know which s satisfy it? Well, we know that if two complex numbers z_1 and z_2 are equal, then both their arguments (angles) are equal ($\arg(z_1) = \arg(z_2)$) and their lengths are equal ($|z_1| = |z_2|$). Let's apply this to equation (4.1.1). If we examine lengths, we find the **length-equation**

$$|G(s)H(s)| = |K|\frac{|s-z_1|\ldots|s-z_m|}{|s-p_1|\ldots|s-p_n|} = |-1| = 1. \tag{4.1.2}$$

This isn't a very useful equation yet. So we also examine the arguments. This time we get the **argument-equation**

$$\arg(K)+\arg(s-z_1)+\ldots+\arg(s-z_m)-\arg(s-p_1)-\ldots-\arg(s-p_n) = \arg(-1) = (2k+1)\cdot 180°, \tag{4.1.3}$$

where $k \in \mathbb{Z}$ can be any constant. (Note that multiplying complex numbers means adding their angle, while dividing complex numbers means subtracting their angle.)

We can even simplify the above equation a bit. We know K must be a real number. And usually it's also positive. This means that $\arg(K) = 0$. So the above equation doesn't even depend on our variable K! And there's even more. If we can find an s satisfying this equation, then we can simply adjust our K, such that it also satisfies the length-equation. So all we need to do is find an s that satisfies the argument-equation. Most of the rules in the upcoming paragraph use this fact.

4.1.2 Root-locus drawing rules

There is no set of rules which will always enable you to draw a perfect root-locus diagram. However, there are rules which can give some info about its shape. So in the following plan of approach, we will mention most of these rules.

1. First we should draw a complex plane. In it, we should draw all the poles and zeroes of $G(s)H(s)$. We denote the n poles by a cross (x) and the m (finite) zeroes by a circle (o). (We will examine the infinite zeroes at point 4.)

2. Next we should examine which points s on the real line are part of the root-locus plot. Just pick a point on this axis. If the total number of poles and zeroes to the right of this point is odd, then this point is part of the root-locus plot. Otherwise, it is not. (By the way, this follows from the argument-equation.) In this way we can find all segments of the real axis which are part of the root-locus plot.

3. We will now find the **breakaway points**. These are the points where the root-locus lines break away from the real axis. To find them, we first have to solve equation (4.1.1) for K. We will find

$$K = -\frac{(s - p_1)(s - p_2)\ldots(s - p_n)}{(s - z_1)(s - z_2)\ldots(s - z_m)}. \tag{4.1.4}$$

We then use the condition $dK/ds = 0$. We will find several solutions s_i for this. However, not all of these solutions represent breakaway points. In fact, to find the actual breakaway points, we need to substitute the values of s back into the above equation. Only if a certain value s_i gives a positive K, then s_i is a valid breakaway point.

4. We can sometimes also find the **asymptotes** of the graph. (However, this only works if $H(s)G(s)$ has enough infinite zeroes. With enough, I mean at least 2. The number of infinite zeroes of $G(s)H(s)$ is equal to $n - m$.) To find the asymptotes, we first have to find the **asymptote center**. It is located at the point

$$s = \frac{p_1 + p_2 + \ldots + p_n - z_1 - z_2 - \ldots - z_m}{n - m}. \tag{4.1.5}$$

From this point, we now draw asymptote lines in $n - m$ directions. These directions have angles

$$\text{asymptote angles} = \frac{\pm(2k + 1) \cdot 180°}{n - m}. \tag{4.1.6}$$

When $|s| \to \infty$ the root-locus plots will converge to these asymptotes.

5. Sometimes we can also find the points where the root-locus plot crosses the imaginary axis. (In fact, this usually only works if $n - m \geq 2$.) On these points, we can write $s = \omega i$. So we can insert this condition into the characteristic equation (4.1.1). We then find a complex equation. If we equate both the real and complex parts of this equation, we can find ω.

6. Finally, we can determine the **angle of departure/angle of arrival** of the complex poles/zeroes. (This, of course, only works if $G(s)H(s)$ has complex poles or zeroes.) We can use the argument-equation for this. Let's suppose we want to find the angle of departure/arrival of the pole p_1. We examine some point s_a, infinitely close to p_1, but at an angle $\theta = \arg(s_a - p_1)$. If we insert this into the argument-equation, we find that

$$\arg(s_a - z_1) + \ldots + \arg(s_a - z_m) - \theta - \arg(s - p_2) - \ldots - \arg(s_a - p_n) = (2k+1) \cdot 180°. \quad (4.1.7)$$

Since s_a is infinitely close to p_1, we can simply substitute $s_a = p_1$ in the above equation. We can then solve for the angle of departure θ at this pole.

Once we have gone through all of these rules, we should be able to draw a more or less accurate graph. In fact, we only have to draw one half of the graph, since root-locus plots are always symmetric about the real axis.

When drawing the root-locus plot, you do have to keep in mind that every line should go from a pole to a (possibly infinite) zero. If you have a zero or a pole without a line, then you probably haven't completed your root-locus plot yet.

4.2   Example problems

The root-locus drawing rules might seem rather vague. So let's clarify them with some examples.

4.2.1   An example problem without infinite zeroes

Let's consider the basic feedback system with $G(s)$ and $H(s)$ given by

$$G(s) = K\frac{(s+2)(s+3)}{s(s+1)} \qquad \text{and} \qquad H(s) = 1. \quad (4.2.1)$$

We want to plot the root-locus graph for this system. In other words, we want the possible range of poles of $F(s)$, if we can vary K. To find this, let's simply apply the rules.

1. The function $G(s)H(s)$ has $n = 2$ poles, being $s = 0$ and $s = -1$. It also has $m = 2$ zeroes, being $s = -2$ and $s = -3$. We draw a complex plane, and insert these four points. (For the final solution, in which these four points are also drawn, see figure 4.2 at the end of this section.)

2. Let's examine a point s on the real axis and count the number of poles/zeroes to the right of it. This number is only odd if $-1 < s < 0$ or $-3 < s < -2$. So only these parts of the real axis are part of our root-locus plot. We mark these sections in our graph. (Again, see figure 4.2.)

3. Let's find the breakaway points. We know that

$$K = -\frac{s(s+1)}{(s+2)(s+3)}. \quad (4.2.2)$$

We can take the derivative with respect to s. This gives us the condition

$$\frac{dK}{ds} = -\frac{(s+2)(s+3)(2s+1) - s(s+1)(2s+5)}{(s+2)^2(s+3)^2} = 0. \quad (4.2.3)$$

Solving this for s gives $s = -\frac{1}{2}(3 \pm \sqrt{3})$. So we have two solutions for s. We do have to check whether they are actual breakaway points. The solution $s = -\frac{1}{2}(3 - \sqrt{3})$ implies that $K = 0.0718$, while $s = -\frac{1}{2}(3 + \sqrt{3})$ gives $K = 13.93$. Both values of K are positive, so both values of s are breakaway points. We can thus draw breakaway points at these positions.

4. Since $n = m$, there are no infinite zeroes. So there are no asymptotes either. This rule therefore isn't of much use.

5. Since there are no asymptotes, there aren't any points where the root-locus plot crosses the imaginary axis.

6. There are no complex poles/zeroes, so we can't find any angles of departure either.

We have got some nice info on the shape of our root-locus diagram already. Experienced root-locus-drawers can already predict its shape. If you are, however, new to root-locus drawing, you might want to find some extra points. We could use the argument-equation for that. We could also simply try to find some point in a brute force kind of fashion. We'll use that second method now, since it's a bit more exact.

We know that root-locus lines always run from a pole to a zero. So there must be some point on the root-locus plot with real part $\text{Re}(s) = -3/2$. What point would that be? To find that out, we can insert $s = -3/2 + \omega i$ into the characteristic equation. If we do this, we can solve for ω. We will find that $\omega = \pm\sqrt{3}/2$. So we just found another point of our root-locus plot!

If you want, you can continue to do this for other points. You will find some more points of the root-locus diagram. When you feel you're sure what the shape of the plot will be, you can draw it. And then you should hopefully have found the plot of figure 4.2.
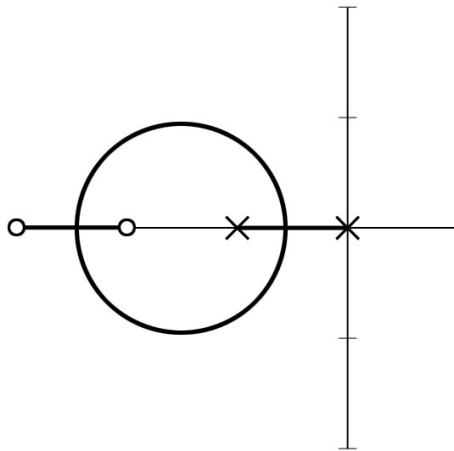

Figure 4.2: The solution of our first example problem.

4.2.2   An example problem with multiple infinite zeroes

This time, let's consider a system with

$$G(s) = \frac{K}{s(s+1)(s^2 + 4s + 13)} \qquad \text{and} \qquad H(s) = 1. \tag{4.2.4}$$

Let's use the drawing rules to draw a root-locus plot.

1. We find that $G(s)H(s)$ has no (finite) zeroes. (So $m = 0$.) It does have four poles. (So $n = 4$.) They are $s = 0$, $s = -1$ and $s = -2 \pm 3i$. We draw those poles in the complex plane.

2. We examine all points on the real axis which have an odd number of poles/zeroes to the right of it. We find that this condition is only valid for $-1 < s < 0$. We mark this part in our plot.

3. Let's find the breakaway points. We have $K = -s(s+1)(s^2 + 4s + 13)$. Taking the derivative will give

$$\frac{dK}{ds} = 4s^3 + 15s^2 + 34s + 13 = 0. \tag{4.2.5}$$

This is a cubic equation. So we let a computer program solve it. We get as solutions $s = -0.467$ and $s = -1.642 \pm 2.067i$. Inserting the first solution in the characteristic equation gives $K = 2.825$, which is positive. So $s = -0.467$ is a breakaway point. We mark this one in our plot. However, inserting the other solutions results in a complex K. They are therefore not breakaway points.

4. We have $n - m = 4$. So there are four infinite zeroes, and thus also four asymptotes. The asymptote center is at

$$s = \frac{z_1 + z_2 + \ldots + z_m - p_1 - p_2 - \ldots - p_n}{n - m} = \frac{-0 - 1 - (2 + \sqrt{3}) - (2 - \sqrt{3})}{4} = -\frac{5}{4}. \tag{4.2.6}$$

The angle of the asymptotes can be found using

$$\text{asymptote angles} = \frac{\pm(2k+1) \cdot 180°}{4}. \tag{4.2.7}$$

If we let k vary, we will find angles $-135°$, $-45°$, $45°$ and $135°$. We lightly mark these asymptotic lines in our plot.

5. There are infinite zeroes. So the root-locus might cross the imaginary axis as well. To find out where it crosses it, we substitute $s = \omega i$ into the characteristic equation. We then find that

$$K = -\omega^4 + 5\omega^3 i + 17\omega^2 - 13\omega i. \tag{4.2.8}$$

If we equate both the real and complex part, we will find that

$$K = -\omega^4 + 17\omega^2 \quad \text{and} \quad 0 = 5\omega^3 - 13\omega. \tag{4.2.9}$$

Solutions for ω are $\omega = 0$ and $\omega = \pm\sqrt{13/5}$. The corresponding solutions for K are $K = 0$ and $K = -(13/5)^2 + 17 \cdot 13/5 = 37.44$, respectively. The solution $\omega = 0$ isn't altogether very surprising. However, we find that the root-locus plot also crosses the imaginary axis at $s = \pm\sqrt{13/5}i = \pm 1.612$. We thus mark those points in our plot.

6. Finally we would like to know the angles of departure of the two complex poles. For that, we use the argument-equation. Let's find the angle of departure θ for the pole $p_1 = -2 + 3i$. First we find the arguments we need to put into the argument-equation. They are

$$\arg((-2 + 3i) - (-2 - 3i)) = \arg(6i) \quad = \quad 90°, \tag{4.2.10}$$
$$\arg((-2 + 3i) - 0) = \arg(-2 + 3i) \quad = \quad 123.69°, \tag{4.2.11}$$
$$\arg((-2 + 3i) - (-1)) = \arg(-1 + 3i) \quad = \quad 108.43°. \tag{4.2.12}$$

It now follows that the angle of departure of p_1 is

$$\theta = (2k+1) \cdot 180° - 90° - 123.69° - 108.43° = -142.13°. \tag{4.2.13}$$

From symmetry we can find that the angle of departure of $p_2 = -2 - 3i$ is $\theta = 142.13°$. We can mark these angles in our root-locus plot.

Now we should have sufficient data to sketch our root-locus plot. The result can be seen in figure 4.3.

4.3 Finding the appropriate points in the plot

It's nice that we know what poles we can give to our system. But how do we choose the right value of K? That's what we'll look at now.
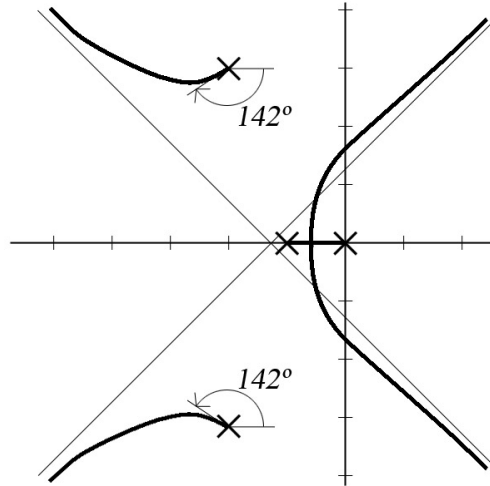
Figure 4.3: The solution of our second example problem.

4.3.1 Setting demands

Let's suppose we have drawn a root-locus plot. How can we use this plot to influence our system? For that, we will examine a pole. But first we have to make sure we examine the right pole of $F(s)$. But which pole is the right pole?

As you know, poles with a bigger real part are more influential. So actually, the pole closest to the origin is the only pole we should focus on. (The other poles are less significant.) Then, using certain **system demands**, we can find this pole.

But what kind of properties can we demand from our system? Let's make a list of that.

- We can set the dampening factor ζ. (Often a value between 0.4 and 0.7 is preferred.) Let's suppose we use this criterion to find a pole p_i. We define the argument of this pole to be $\phi = \arg(p_i)$. It can then be shown that $\zeta = \cos(180° - \phi)$.

  So how do we find p_i? Well, we draw a line from the origin with an angle $\phi = 180° - \arccos\zeta$. The point where it intersects our root-locus plot is the position of our pole p_i.

- We can also set the damped natural frequency ω_n. It can now be shown that the undamped natural frequency due to some pole p_i is $|p_i| = \omega_n$.

  So how do we find p_i? We simply draw a circle with radius ω_n around the origin. The point where it intersects the root-locus plot is where our pole p_i is.

- Finally we can set the undamped natural frequency ω_d. We can find that the damped natural frequency due to a pole p_i is equal to the real part of p_i. So, $\mathrm{Re}(p_i) = \pm\omega_d$.

  So how do we find p_i? We simply draw a horizontal line at a height $\pm\omega_d$. The point where it intersects the root-locus plot is where our pole p_i is.

4.3.2 Finding the gain K

We now know how to find the right pole. But how do we use this to find K? Well, we can simply use the characteristic equation for this. Insert the pole p_i we just found into the characteristic equation. From this, you can solve for K quite easily. And once you know K, all that is left to do is to find the other poles of $F(s)$.

5. Analyzing frequency response

We have previously seen how to analyze systems subject to basic input functions. We have even learned how to design them. Now we'll examine systems subject to periodic functions.

5.1 Bode diagrams

The behaviour of a system subject to periodic input is called the **frequency response**. We want to examine this. A very nice way to do this, is by making Bode diagrams. In this part we'll examine how to draw these diagrams.

5.1.1 Subjecting systems to a sine function input

Let's suppose we have a system with transfer function $F(s)$. We give it as input a sine function. So,

$$u(t) = X \sin \omega t, \qquad \text{or, equivalently} \qquad U(s) = \frac{\omega X}{s^2 + \omega^2}. \tag{5.1.1}$$

Here X is the **input amplitude** and ω is the **frequency**. We want to know how this input effects the output $y(t)$. To find that, we first examine $Y(s)$. Using a partial-fraction expansion, we can show that

$$Y(s) = F(s)U(s) = \frac{a}{s + \omega i} + \frac{\bar{a}}{s - \omega i} + F(s), \qquad \text{where} \quad a = -\frac{XF(-\omega i)}{2i} \quad \text{and} \quad \bar{a} = \frac{XF(\omega i)}{2i}. \tag{5.1.2}$$

We assume that our system is a stable system. In this case the effects due to initial conditions eventually fade out. The so-called **steady-state solution** for $y(t)$ (noted as $y_{ss}(t)$) is then given by

$$y_{ss}(t) = X|F(\omega i)| \sin(\omega t + \phi) = Y \sin(\omega t + \phi), \qquad \text{where} \qquad \phi = \arg(F(\omega i)). \tag{5.1.3}$$

Here Y is the **output amplitude**. Also ϕ is the **phase angle**.

A few important things can be noted. We note that the frequency of the output is the same as the frequency of the input. However, the amplitude is not the same. It is magnified by an **amplitude gain** $K = Y/X = |F(\omega i)|$. There is also a phase difference, having magnitude $\phi = \arg(F(\omega i))$. By the way, if $\phi < 0$, then there is so-called **phase lag**. The system/network is then called a **lag network**. Similarly, if $\phi > 0$, then there is **phase lead**. The network is then called a **lead network**.

5.1.2 Introduction to Bode diagrams

We want to know how systems respond to periodic inputs. So it would be nice to know how the amplitude gain K and the phase angle ϕ depend on the frequency ω. We display that in so-called **Bode diagrams**. These diagrams actually consist of two diagrams.

In the first diagram (the **gain Bode diagram**) we display the gain K versus the frequency ω. However, both values can differ rather much. So to get some meaningful graphs, we don't just display the gain K. Instead, we first transform this gain K to the unit dB. We do this using

$$K_{dB} = 20 \log K. \tag{5.1.4}$$

This **decibel gain** is then plotted versus $\log \omega$. The plot is therefore a **log-log plot**. (By the way, the base of the logarithm in both cases is 10.)

In the second diagram (the **phase angle Bode diagram**) we display the phase angle ϕ versus the logarithm of the frequency. So we plot ϕ versus $\log \omega$.

Since both parts of the Bode diagram have the same horizontal axis, they are usually placed above each other. This causes the decibel gain and the phase angle, corresponding to the same frequency, to be above each other as well.

One important question remains. How do we find these diagrams? Well, we simply insert $s = \omega i$ into the transfer function $F(s)$. This gives us a complex number. We then determine the length and argument of this number. These are then the gain K and phase angle ϕ. To investigate how this works, we will now examine some basic transfer functions $F(s)$ and describe their Bode diagrams.

5.1.3   A constant transfer function

If $F(s)$ is constant ($F(s) = K$), then the gain is simply K. In other words, the gain Bode diagram is simply a horizontal line, situated at $K_{dB} = 20 \log K$.

Usually this gain K is a positive number. The argument of a positive number is simply $\phi = 0$. So the phase angle Bode diagram of $F(s)$ is simply a horizontal line at $\phi = 0$.

5.1.4   A rather simple transfer function

Now let's suppose $F(s) = 1/s$. In this case $F(\omega i) = 1/\omega i = -i/\omega$. The decibel gain now is

$$K_{dB} = 20 \log |-i/\omega| = 20 \log 1/\omega = -20 \log \omega. \tag{5.1.5}$$

Remember that we plotted K_{dB} versus $\log \omega$. This means that the gain Bode diagram of $F(s)$ is simply a straight line. It has a slope of -20. (So the line is directed downward.) Also, if $\omega = 1$, then $\log \omega = 0$ and thus also $K_{dB} = 0$.

We also examine the phase angle Bode diagram. This one is rather easy. We can find that $\phi = \arg(-i/\omega) = -90°$. The graph is thus a horizontal line.

5.1.5   A first order transfer function

Now we examine $F(s) = 1/(1 + Ts)$. This gives us $F(\omega i) = 1/(1 + T\omega i)$. The decibel gain now becomes

$$K_{dB} = 20 \log \left| \frac{1}{1 + T\omega i} \right| = -20 \log \sqrt{1 + \omega^2 T^2}. \tag{5.1.6}$$

Let's investigate the factor ωT more closely. If ωT is very small, then $K_{dB} \approx 0$. So for small ωT, we simply have a horizontal line positioned at $K_{dB} = 0$. If, however, ωT is very big, then $K_{dB} \approx -20 \log \omega T$. It can be shown that this is a straight line with slope -20. (It thus goes downward.) It crosses the $0dB$ axis at $\omega = 1/T$. This frequency is called the **corner frequency** or **break frequency**.

We can actually apply a trick now. If we want to draw the gain Bode diagram of $F(s)$, then we only have to draw the two straight lines we just described. This gives a good approximation of the actual Bode diagram already. The only place where it is seriously off is near the corner frequency $\omega = 1/T$. (It is off by about $3dB$ there.) If we take into account for this deviation, then we have actually drawn a rather accurate curve.

But what about the phase angle Bode diagram? Well, we can find that the phase angle is

$$\phi = -\arctan \omega T. \tag{5.1.7}$$

But what does this mean? For small ωT, the value of ϕ converges to $0°$. For big ωT, we have $\phi \approx -90°$. Finally at the corner frequency $\omega = 1/T$ we have $\phi = -45°$. Just draw a smooth curve through these points and you've drawn the phase angle Bode diagram.

5.1.6   A second order transfer function

Finally, we examine $F(\omega i) = 1/(1 + 2\zeta i\omega/\omega_n + (i\omega/\omega_n)^2)$. The decibel gain is now given by

$$-20\log\sqrt{\left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}. \tag{5.1.8}$$

We will examine ω/ω_n. If this quantity is very small, then we have $K_{dB} \approx 0$. So this is a horizontal line at K_{dB}. (Does it sound familiar?) If ω/ω_n is very big, then $K_{dB} \approx -40\log\omega/\omega_n$. In other words, we find a straight line again. However, the slope of this straight line now is -40.

But what happens if $\omega \approx \omega_n$? Now something interesting occurs. The gain Bode diagram actually goes up. If the input frequency is close to the natural frequency, the amplitude gain increases. There is thus a peak in the gain Bode diagram. This phenomenon is called **resonance**. The **resonant frequency** ω_r is defined as the frequency at which the gain K is maximal. This frequency is given by

$$\omega_r = \omega_n\sqrt{1 - 2\zeta^2}. \tag{5.1.9}$$

So resonance only occurs if $\zeta < \sqrt{2}/2$. By the way, the corresponding value of the gain K is called the **resonant peak value** M_r. It is given by

$$M_r = |G(\omega i)|_{max} = |G(\omega_r i)| = \frac{1}{2\zeta\sqrt{1 - \zeta^2}}. \tag{5.1.10}$$

So as $\zeta \to 0$, then $M_r \to \infty$.

Well, so much for the gain Bode diagram. Now let's examine the phase angle Bode diagram. This time ϕ is given by

$$\phi = -\arctan\left(\frac{2\zeta\omega/\omega_n}{1 - (\omega/\omega_n)^2}\right). \tag{5.1.11}$$

What can we derive from this? For small ω/ω_n, we have $\phi \approx 0°$. For big ω/ω_n we have $\phi \approx -180°$. If $\omega = \omega_n$, then $\phi = -90°$.

Surprisingly, none of these points depend on ζ. You may wonder, doesn't the shape of this graph depend on ζ at all? Well, ζ does have effect. It determines the 'smoothness' of the graph. If $\zeta \approx 1$, then the graph of ϕ goes from $0°$ to $-180°$ rather slowly. If, however, $\zeta \approx 0.1$, then this transition occurs rather sudden.

5.1.7   Powers of basic functions

We now know what the Bode diagrams of basic transfer function $F(s)$ looks like. But what if we get a power of such a basic function? For example, let's suppose that $F(s) = 1/(1 + Ts)$. What should we do if we need to make a Bode diagram of $F(s)^n = 1/(1 + Ts)^n$?

Well, there is a trick we can use. We know that the decibel gain of $F(s)$ is $K_{dB} = 20\log|F(\omega i)|$. The decibel gain of $F(s)^n$ then becomes

$$K_{dB} = 20\log|F(\omega i)^n| = 20\log|F(\omega i)|^n = 20n\log|F(\omega i)|. \tag{5.1.12}$$

So the decibel gain simply becomes n times as big! But what happens to the phase angle? Well, the phase angle of $F(s)^n$ is equal to the argument of $F(\omega i)^n$. ($\phi = \arg(F(\omega i)^n)$.) And there is a rule in complex number theory that states that $\arg(\lambda^n) = n\arg(\lambda)$. This means that ϕ also becomes n times as big!

What can we conclude from this? If we know the Bode diagram of a certain function $F(s)$, then finding the diagram of $F(s)^n$ is just a matter of multiplication.

By the way, this also works if n is negative. For example, let's suppose we want to find the Bode diagram of $F(s) = 1 + \omega T$. Then we just multiply the Bode diagram of $F(s) = 1/(s + \omega T)$ by -1. (In other words, mirror it about the horizontal axis.)

5.1.8   Bode diagrams of general functions

Now let's suppose we have some transfer function $F(s)$ that can be written as

$$F(s) = K\frac{(1 + T_{z1}s)(1 + T_{z2}s)\ldots(1 + T_{zm}s)}{(1 + T_{p1}s)(1 + T_{p2}s)\ldots(1 + T_{pn}s)}, \tag{5.1.13}$$

how can we draw the Bode diagram? Again, we can use a trick. And it is similar to the trick of the previous paragraph.

There is a rule stating $\log ab = \log a + \log b$. There is another rule stating that $\arg(ab) = \arg(a) + \arg(b)$. What can we derive from this? Well, if we want to find the Bode diagram of $F(s)$, we simply look at all the separate Bode diagrams. So we examine the bode diagrams of $K, (1 + T_{z1}s), \ldots (1 + T_{zm}s), (1 + T_{p1}s), \ldots, (1 + T_{pn}s)$. Then (according to the rules above) we can simply add up all these Bode diagrams. This then gives us the Bode diagram of $F(s)$, which is exactly what we want.

5.1.9   Cutoff parameters

So what have we learned thus far? We have seen that a system can pass on certain frequencies to its output. However, other frequencies are damped out. A Bode diagram clearly displays which frequencies are damped out. However, we can also express this in certain parameters.

First we define the **zero-frequency gain** K_{dB_0}. This is the decibel gain K_{dB} as $\omega \to 0$. The **cutoff frequency** ω_b now satisfies

$$K_{dB}(\omega) < K_{dB_0} - 3dB \qquad \text{for every } \omega > \omega_b. \tag{5.1.14}$$

In other words, the system can't really transmit any frequency higher than the cutoff frequency. However, the system can transmit frequencies between 0 and ω_b. This frequency range is therefore called the **bandwidth** of the system.

Finally, we are also interested in how fast the decibel gain K_{dB} drops below $K_{dB_0} - 3dB$. In other words, we want to know the slope of the gain Bode diagram at $\omega = \omega_b$. This slope is called the **cutoff rate**. A high cutoff rate usually indicates a high resonance peak, and therefore a less stable system.

5.2   Nyquist plots and Nyquist stability

Another way of examining stability is by using the Nyquist method. We will delve into that method in this part.

5.2.1   Nyquist plots

Nyquist plots (also known as **polar plots**) are rather different from Bode diagrams. In a Nyquist plot the gain K and the phase angle ϕ are plotted in polar coordinates. The gain K then is the distance from the origin, and ϕ is the angle.

Let's take a closer look at what we're actually plotting. We have as radius $K = |F(\omega i)|$ and as angle $\phi = \arg(F(\omega i))$. So, in fact, what we plot is the function $F(\omega i)$ itself! If we let ω vary from 0 to ∞ and plot $F(\omega i)$ in the complex plane, then we will find the Nyquist plot!

Nyquist plots are often very fun. There is, however, one slight problem with them. The trick of adding up basic solutions (which worked so well for Bode diagrams) doesn't apply here. So, instead, we just have to draw these diagrams in the old fashioned way: calculate a few points (for certain ω) and draw a nice line through them.

5.2.2 The mapping theorem

Soon we will delve into the Nyquist stability criterion. But before we do that, we will look at the **mapping theorem**. (We will need it for the Nyquist stability criterion.)

Suppose we have some complex function $L(s)$, which is the ratio of two polynomials. We now take a closed contour in the complex plane. We say that there are P poles within this contour and Z zeroes. (There may not be any poles or zeroes on the contour itself.)

For every value s on this contour, we will now map the value of $L(s)$. This gives as a second closed curve. We now walk in a clockwise direction along this second curve. We call N the amount of times we have walked around the origin of the complex plane. The mapping theorem now states that $N = Z - P$.

5.2.3 The Nyquist stability criterion

It's finally time to examine the Nyquist stability criterion. This criterion states something about the stability of a basic closed-loop system.

Let's suppose we have a basic closed-loop system. The transfer function of this system is $F(s) = G(s)/(1 + G(s)H(s))$. If the system is stable, then all the poles of $F(s)$ must lie in the left half of the complex plane. We can also see that the poles of $F(s)$ are the roots of the denominator $1 + G(s)H(s)$. So the roots of $L(s) = 1 + G(s)H(s)$ must also lie in the left half of the complex plan.

Let's examine this function $L(s)$ more closely. In fact, let's apply the mapping theorem to $L(s)$. First we will define the first curve. And now comes the brilliant step: we simply take the curve which encloses the right half of the complex plane. Then P is the number of zeroes in the right half of the complex plane. Similarly, Z is the number of zeroes.

Let's examine this curve more closely. One part of this curve consists of the imaginary axis. However, the other part has a distance infinitely far away from the origin. That seems complicated. However, if $\lim_{s \to \infty} L(s) = $ constant, then it can be shown that this second part isn't interesting at all. So, we actually only need to concern ourselves with the imaginary axis!

Now let's draw the second curve. Our second curve is formed by $L(s)$, where s can be any point on the imaginary axis. In other words, our second curve is formed by $L(\omega i)$, where ω can be any real value. So, this second curve is simply the Nyquist plot of $L(s)$! Well, it almost is. In a Nyquist plot ω ranged from 0 to ∞. But now it goes from $-\infty$ to ∞. This isn't really a problem. Just take the Nyquist plot of $L(s)$, and mirror it with respect to the x-axis. That will give you the complete plot for $-\infty < \omega < \infty$.

Let's suppose we have found the entire second curve. To find N, we can simply count the number of times which the curve goes around the origin. But what can we conclude from this? Well, we know that $N = Z - P$. To ensure stability, there may not be any zeroes of $L(s)$ in the right half of the complex plane. We therefore must have $Z = 0$. So, the system is stable if, and only if, $N = -P$. Usually the poles of $L(s)$ (or equivalently, the poles of $G(s)H(s)$) are known. So we can readily check whether the condition $N = -P$ holds.

5.2.4 Sidenotes to the Nyquist stability criterion

There is only one slight problem with the Nyquist stability criterion. We have applied the mapping theorem. And the mapping theorem demanded that no poles/zeroes lie on the first curve. In other

words, no curves/zeroes of $L(s)$ may lie on the imaginary axis. This may sound like a problem, but luckily there is a solution. We simply change our curve slightly. If we encounter a pole/zero on the imaginary axis, we simply go around it. In other words, we take an infinitely small half circle, and place it around the pole/zero. This then solves our problem.

By the way, we also may alter the above criterion slightly. Previously we have plotted $L(\omega i) = 1 + G(\omega i)H(\omega i)$ and counted how many times the curve went around the origin. Now let's do something which is only slightly different. Instead, we plot $G(\omega i)H(\omega i)$. This time we count how many times the curve goes around the point -1. This small trick doesn't change the number N, so we may readily apply it.

5.2.5 Crossover frequencies and phase/gain margins

Let's suppose we have the Nyquist plot of $G(s)$. According to the Nyquist stability criterion, we can determine whether it is stable. We can also deduce something else from this criterion. Namely that, if the Nyquist plot goes through the point -1, then the system is on the verge of instability. In this case there is some frequency ω for which $|G(\omega i)| = 1$ and $\arg(G(\omega i)) = 180°$.

How can we check whether this is the case? For that, we first find the gain crossover frequency and the phase crossover frequency.

- The **gain crossover frequency** ω_{gc} is the frequency for which the gain is equal to $|G(\omega i)| = 1$.

- The **phase crossover frequency** ω_{pc} is the frequency for which the phase angle is $\arg(G(\omega i)) = 180°$.

(Note that there can be multiple gain/phase crossover frequencies.) If any of the gain crossover frequencies is equal to any of the phase crossover frequencies, then we know that the system is on the verge of instability.

Of course $G(\omega i)$ doesn't always pass through the point -1. But even if it doesn't, we would still like to know how close $G(\omega i)$ is to this point. (Since, if the curve goes closer to -1, the system becomes less stable.) To examine this 'closeness', we define the phase and gain margin.

- Let's examine the angle $\phi = \arg(G(\omega_{gc} i))$ at the gain crossover frequency. The **phase margin** γ now is the difference between this angle and the angle of -1 (which is $180°$). So $\gamma = \phi - 180°$.

- Let's examine the gain $|G(\omega_{pc} i)|$ at the phase crossover frequency. The **gain margin** K_g is the reciprocal of this magnitude. So, $K_g = 1/|G(\omega_{pc} i)|$. (In decibels this becomes $K_{g_{dB}} = -20 \log |G(\omega_{pc} i)|$.)

For most systems, both margins should be positive to ensure stability. Negative margins are often signs of instability. Although this does depend on the system itself.

5.3 Nichols plots

Let's suppose we have a simple closed loop system. The closed-loop transfer function is given by $F(s) = G(s)/(1 + G(s))$. We also know the frequency response of the open-loop transfer function $G(s)$. Can we derive from this the frequency response of $F(s)$ itself? It turns out that we can. We will find out how in this part.

5.3.1 M circles and N circles

First we make some definitions. Let's define the closed loop gain $M = |F(\omega i)|$ and the closed loop phase angle $\alpha = \arg(F(\omega i))$. We also define $N = \tan \alpha$. And finally we write $G(\omega j)$ as $G(\omega j) = X + Yi$.

Now let's look at the magnitude of $F(\omega i)$. This gives us an expression for M, being

$$M = |F(\omega i)| = \frac{|G(s)|}{1 + |G(s)|} = \frac{|X + Yi|}{|1 + X + Yi|} = \frac{\sqrt{X^2 + Y^2}}{\sqrt{(1 + X)^2 + Y^2}}. \tag{5.3.1}$$

If $M = 1$, then we find that $X = -1/2$ (for every Y). So if we make a plot of Y versus X, we will get a vertical line. If $M \neq 1$, then we can rewrite the above equation to

$$\left(X + \frac{M^2}{M^2 - 1}\right)^2 + Y^2 = \left(\frac{M}{M^2 - 1}\right)^2. \tag{5.3.2}$$

But this is the equation of a circle! The circle has a radius $|M/(M^2 - 1)|$! The center of the circle is located at $X = -M^2/(M^2 - 1)$ and $Y = 0$. The circles we can find with these equation are called M **circles** or **constant magnitude loci**. Along these circles, the magnitude of $F(\omega i)$ is constant.

Similarly, we can find an expression for α (or equivalently, for N), by looking at the argument of $F(\omega i)$. This time we can derive that

$$\left(X + \frac{1}{2}\right)^2 + \left(Y - \frac{1}{2N}\right)^2 = \frac{1}{4} + \left(\frac{1}{2N}\right)^2. \tag{5.3.3}$$

This is also the expression for a circle! This time the radius is $\sqrt{1/4 + 1/(2N)^2}$. The center of the circle lies at $X = -1/2$ and $Y = 1/(2N)$. The circles we can find with this equation are called N **circles** or **constant phase angle loci**. Along these circles, the argument of $F(\omega i)$ is constant.

So how do we derive the frequency response of $F(s)$ from the frequency response of $G(s)$? Well, we simply make a Nyquist plot of $G(s)$. We then lay this plot over the plots of the M circles and N circles. From this we can derive the value of M and N (and thus α) for every frequency ω.

5.3.2 Nichols charts

Working with multiple graphs (as was done in the previous paragraph) may be a bit inconvenient. That's why often **Nichols charts** are used. The horizontal axis of such a Nichols chart is the phase angle ϕ of $G(s)$. The vertical axis is the decibel gain K_{dB} of $G(s)$. In this graph, both the M circles and N circles are plotted. (Although they aren't exactly circles anymore.)

So how do we use such a Nichols chart? Once more we plot the frequency reponse properties of $G(s)$: We plot the decibel gain K_{dB} versus the phase angle ϕ. We then lay this curve over a Nichols chart. We mark the points where the curve intersects the M and N circles. These points then indicate the corresponding values of M and N (and thus of α). By taking enough intersection points, we can actually derive the frequency response of $F(s)$ quite accurately.