Chapter 7

Loop Analysis of Feedback Systems

This lecture describes how to analyze the stability and robustness of a feedback system by looking at the open loop transfer function. We introduce the Nyquist criteria for stability and talk about the gain and phase margin as measures of robustness.

7.1 Introduction

The basic idea of loop analysis is to "trace" a signal around a feedback loop to determine whether the signal grows or decays when it is fed back on itself through the system dynamics. It will be convenient to make use of a different way of plotting the frequency response of the system with transfer function $G(s)$. The frequency response can be represented in the complex plane by graphically by plotting the magnitude and phase of $G(i\omega)$ for all frequencies, as shown in Figure 7.1. Such a graph is called the Nyquist plot. The magnitude $a = |G(i\omega)|$ represents the gain and the angle $\phi = \arg G(i\omega)$ represents the phase shift. The phase shift is typically negative which implies that the output will lag the input.

The Nyquist plot gives us a way of looking at the stability of a feedback system. Consider the feedback system in Figure 6.17 consisting of a process with transfer function $P(s)$ and a controller with transfer function $C(s)$. Introduce the loop transfer function

$$L(s) = P(s)C(s), \tag{7.1}$$

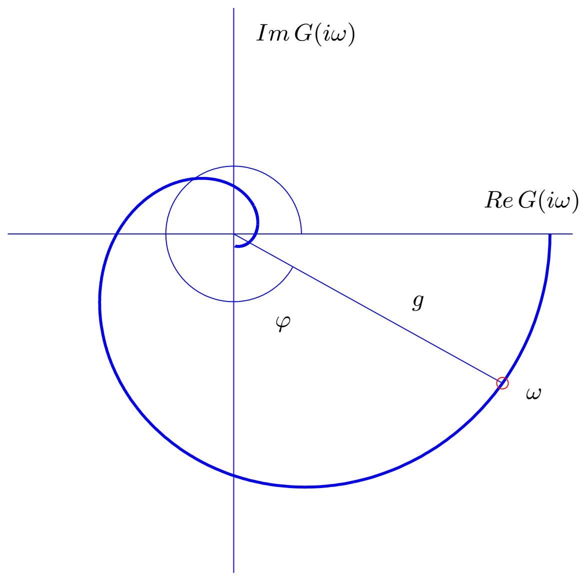

Figure 7.1: Nyquist plot of the transfer function $G(s) = \dfrac{1.4e^{-s}}{(s+1)^2}$. The gain and phase for the frequency ω are $g = |G(i\omega)|$ and $\varphi = \arg G(i\omega)$.

which represents signal transmission around the loop. The system can then be represented by the block diagram in Figure 7.2.

We will first determine conditions for having a periodic oscillation in the

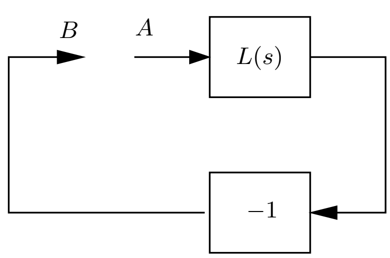

Figure 7.2: Block diagram of a simple feedback system.

loop. Assume that the feedback loop is cut as indicated in the figure and that a sinusoid of frequency ω is injected at point A. In steady state the signal at point B will also be a sinusoid with the same frequency. It seems reasonable that an oscillation can be maintained if the signal at B has the same amplitude and phase as the injected signal because we could then connect A to B. Tracing signals around the loop we find that the condition that the signal at B is identical to the signal at A is that

$$L(i\omega_0) = -1 \tag{7.2}$$

This condition has a nice interpretation in the Nyquist plot. It means that the Nyquist plot of $L(i\omega)$ intersects the negative real axis at the point -1. The frequency where the intersection occurs is the frequency of the oscillation.

Intuitively it seems reasonable that the system would be stable if the critical point -1 is on the left hand side of the Nyquist curve as indicated in Figure 7.2. This means that the signal at point B will have smaller amplitude than the injected signal. This is essentially true, but there are several subtleties, that requires a proper mathematics to clear up. This will be done later. The precise statement is given by Nyquist's stability criterion.

7.2 Nyquist Criterion

We will now state and prove the Nyquist stability theorem. This will require more results from the theory of complex variables than in many other parts of the book. Since precision is needed we will also use a more mathematical style of presentation. The key result is the following theorem about functions of complex variables.

Theorem 7 (Principle of Variation of the Argument). Let D be a closed region in the complex plane and let Γ be the boundary of the region. Assume the function f is analytic in D and on Γ except at a finite number of poles and zeros, then

$$w_n = \frac{1}{2\pi}\Delta_\Gamma \arg f(z) = \frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}dz = N - P$$

where N is the number of zeros and P the number of poles in D. Poles and zeros of multiplicity m are counted m times. The number w_n is called the winding number and $\Delta_\Gamma \arg f(z)$ is the variation of the argument of the function f as the curve Γ is traversed in the positive direction.

Proof. Assume that $z = a$ is a zero of multiplicity m. In the neighborhood of $z = a$ we have

$$f(z) = (z - a)^m g(z)$$

where the function g is analytic and different form zero. We have

$$\frac{f'(z)}{f(z)} = \frac{m}{z - a} + \frac{g'(z)}{g(z)}$$

The second term is analytic at $z = a$. The function f'/f thus has a single pole at $z = a$ with the residue m. The sum of the residues at the zeros of the function is N. Similarly we find that the sum of the residues of the poles of is $-P$. Furthermore we have

$$\frac{d}{dz} \log f(z) = \frac{f'(z)}{f(z)}$$

which implies that

$$\int_\Gamma \frac{f'(z)}{f(z)} dz = \Delta_\Gamma \log f(z)$$

where Δ_Γ denotes the variation along the contour Γ. We have

$$\log f(z) = \log |f(z)| + i \arg f(z)$$

Since the variation of $|f(z)|$ around a closed contour is zero we have

$$\Delta_\Gamma \log f(z) = i \Delta_\Gamma \arg f(z)$$

and the theorem is proven. □

Remark 7. The number w_n is called the winding number.

Remark 8. The theorem is useful to determine the number of poles and zeros of an function of complex variables in a given region. To use the result we must determine the winding number. One way to do this is to investigate how the curve Γ is transformed under the map f. The variation of the argument is the number of times the map of Γ winds around the origin in the f-plane. This explains why the variation of the argument is also called the winding number.

We will now use the Theorem 7 to prove Nyquist's stability theorem. For that purpose we introduce a contour that encloses the right half plane. For that purpose we choose the contour shown in Figure 7.3. The contour consists of a small half circle to the right of the origin, the imaginary axis and

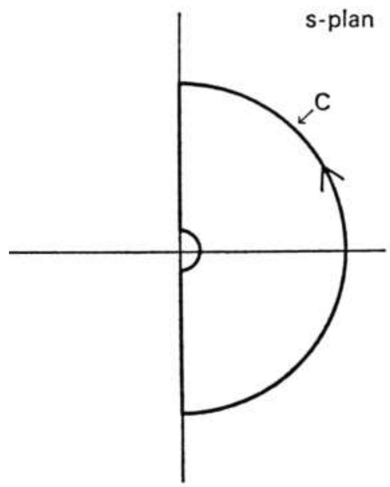

Figure 7.3: Contour Γ used to prove Nyquist's stability theorem.

a large half circle to the right with with the imaginary axis as a diameter. To illustrate the contour we have shown it drawn with a small radius r and a large radius R. The Nyquist curve is normally the map of the positive imaginary axis. We call the contour Γ the full Nyquist contour.

Consider a closed loop system with the loop transfer function $L(s)$. The closed loop poles are the zeros of the function

$$f(s) = 1 + L(s)$$

To find the number of zeros in the right half plane we thus have to investigate the winding number of the function $f = 1 + L$ as s moves along the contour Γ. The winding number can conveniently be determined from the Nyquist plot. A direct application of the Theorem 7 gives the following.

Theorem 8 *(Nyquist's Stability Theorem). Consider a simple closed loop system with the loop transfer function $L(s)$. Assume that the loop transfer function does not have any poles in the region enclosed by Γ and that the winding number of the function $1 + L(s)$ is zero. Then the closed loop characteristic equation has not zeros in the right half plane.*

We illustrate Nyquist's theorem by an examples.

Example 26 *(A Simple Case).* Consider a closed loop system with the loop transfer function

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

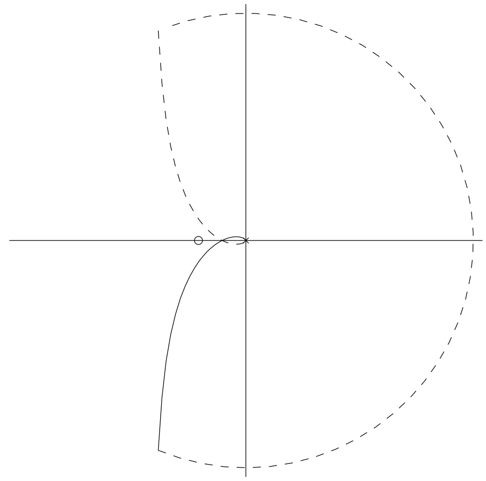

Figure 7.4: Map of the contour Γ under the map $L(s) = \frac{k}{s((s+1)^2}$. The curve is drawn for $k < 2$. The map of the positive imaginary axis is shown in full lines, the map of the negative imaginary axis and the small semi circle at the origin in dashed lines.

Figure 7.4 shows the image of the contour Γ under the map L. The Nyquist plot intersects the imaginary axis for $\omega = 1$ the intersection is at $-k/2$. It follows from Figure 7.4 that the winding number is zero if $k < 2$ and 2 if $k > 2$. We can thus conclude that the closed loop system is stable if $k < 2$ and that the closed loop system has two roots in the right half plane if $k > 2$.

By using Nyquist's theorem it was possible to resolve a problem that had puzzled the engineers working with feedback amplifiers. The following quote by Nyquist gives an interesting perspective.

> Mr. Black proposed a negative feedback repeater and proved by tests that it possessed the advantages which he had predicted for it. In particular, its gain was constant to a high degree, and it was linear enough so that spurious signals caused by the interaction of the various channels could be kept within permissible limits. For best results, the feedback factor, the quantity usually known as $\mu\beta$ (the loop transfer function), had to be numerically much larger than unity. The possibility of stability with a feedback factor greater than unity was puzzling. Granted that the factor is negative it was not obvious how that would help. If the factor was -10, the effect of one round trip around the feedback loop is to change the magnitude of the current from, say 1 to -10. After a second trip around the loop the current becomes 100, and

> so forth. The totality looks much like a divergent series and it was not clear how such a succession of ever-increasing components could add to something finite and so stable as experience had shown. The missing part in this argument is that the numbers that describe the successive components 1, -10, 100, and so on, represent the steady state, whereas at any finite time many of the components have not yet reached steady state and some of them, which are destined to become very large, have barely reached perceptible magnitude. My calculations were principally concerned with replacing the indefinite diverging series referred to by a series which gives the actual value attained at a specific time t. The series thus obtained is convergent instead of divergent and, moreover, converges to values in agreement with the experimental findings.
>
> This explains how I came to undertake the work. It should perhaps be explained also how it come to be so detailed. In the course of the calculations, the facts with which the term conditional stability have come to be associated, became apparent. One aspect of this was that it is possible to have a feedback loop which is stable and can be made unstable by by increasing the loop loss. this seemed a very surprising result and appeared to require that all the steps be examined and set forth in full detail.

This quote clearly illustrate the difficulty in understanding feedback by simple qualitative reasoning. We will illustrate the issue of conditional stability by an example.

Notice that Nyquist's theorem does not hold if the loop transfer function has a pole in the right half plane. There are extensions of the Nyquist theorem to cover this case but it is simpler to invoke Theorem **??** directly. We illustrate this by two examples.

Example 27 *(Loop Transfer Function with RHP Pole).* Consider a feedback system with the loop transfer function

$$L(s) = \frac{k}{s(s-1)(s+5)}$$

This transfer function has a pole at $s = 1$ in the right half plane. This violates one of the assumptions for Nyquist's theorem to be valid. The Nyquist plot of the loop transfer function is shown in Figure 7.5. Traversing the contour Γ in clockwise we find that the winding number is 1. Applying Theorem 1 we find that

$$N - P = 1$$

Since the loop transfer function has a pole in the right half plane we have $P = 1$ and we get $N = 2$. The characteristic equation thus has two roots in the right half plane.

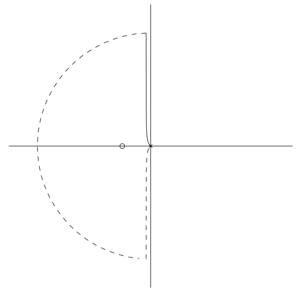

Figure 7.5: Map of the contour Γ under the map $L(s) = \frac{k}{s(s-1)(s+5)}$. The curve on the right shows the region around the origin in larger scale. The map of the positive imaginary axis is shown in full lines, the map of the negative imaginary axis and the small semi circle at the origin in dashed lines.

Example 28 *(The Inverted Pendulum).* Consider a closed loop system for stabilization of an inverted pendulum with a PD controller. The loop transfer function is

$$L(s) = \frac{s+2}{s^2-1} \tag{7.3}$$

This transfer function has one pole at $s = 1$ in the right half plane. The Nyquist plot of the loop transfer function is shown in Figure 7.6. Traversing the contour Γ in clockwise we find that the winding number is -1. Applying Theorem 1 we find that

$$N - P = -1$$

Since the loop transfer function has a pole in the right half plane we have $P = 1$ and we get $N = 0$. The characteristic equation thus has no roots in the right half plane and the closed loop system is stable.

7.3 Small Gain Theorem

7.4 Stability Margins

In practice it is not enough that the system is stable. There must also be some margins of stability. There are many ways to express this. Many of the criteria are inspired by Nyquist's stability criterion. They are based on the fact that it is easy to see the effects of changes of the gain and the phase of the controller in the Nyquist diagram of the loop transfer function $L(s)$.

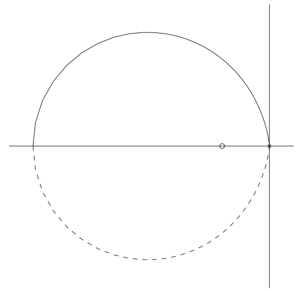

Figure 7.6: Map of the contour Γ under the map $L(s) = \frac{s+2}{s^2-1}$ given by (7.3). The map of the positive imaginary axis is shown in full lines, the map of the negative imaginary axis and the small semi circle at the origin in dashed lines.

An increase of controller gain simply expands the Nyquist plot radially. An increase of the phase of the controller twists the Nyquist plot clockwise, see Figure 7.7. The gain margin g_m tells how much the controller gain can be increased before reaching the stability limit. Let ω_{180} be the smallest frequency where the phase lag of the loop transfer function $L(s)$ is 180°. The gain margin is defined as

$$g_m = \frac{1}{|L(i\omega_{180})|} \tag{7.4}$$

The phase margin φ_m is the amount of phase lag required to reach the stability limit. Let ω_{gc} denote the lowest frequency where the loop transfer function $L(s)$ has unit magnitude. The phase margin is then given by

$$\varphi_m = \pi + \arg L(i\omega_{gc}) \tag{7.5}$$

The margins have simple geometric interpretations in the Nyquist diagram of the loop transfer function as is shown in Figure 7.7.

A drawback with gain and phase margins is that it is necessary to give both numbers in order to guarantee that the Nyquist curve not is close to the critical point. One way to express margins by a single number is to use the shortest distance from the Nyquist curve to the critical point. We call this number the stability margin. This number also has other nice interpretations as will be discussed in Chapter **??**.

Reasonable values of the margins are phase margin $\varphi_m = 30° - -60°$, gain margin $g_m = 2 - 5$, and shortest distance to the critical point $d = 0.5 - 0.8$.

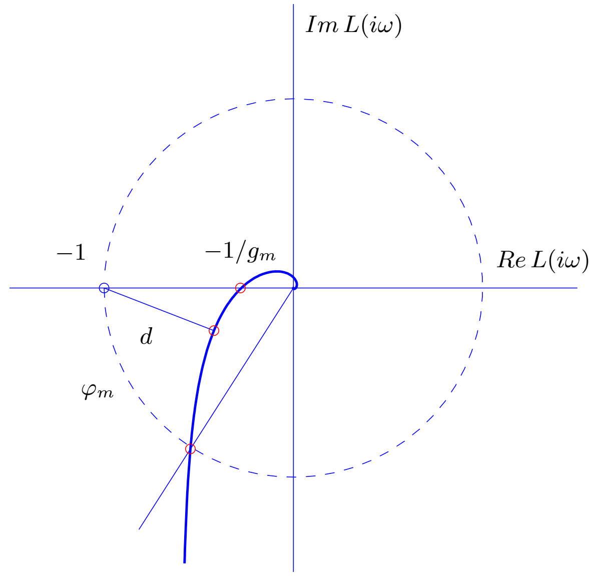

Figure 7.7: Nyquist plot of the loop transfer function L with gain margin g_m, phase margin φ_m and stability margin d.

The gain and phase margins were originally used for the case when the Nyquist plot intersects the unit circle and the negative real axis once. For more complicated systems there may be many intersections and it is then necessary to consider the intersections that are closest to the critical point. For more complicated systems there is also another number that is highly relevant namely the delay margin. The delay margin is defined as the smallest time delay required to make the system unstable. For loop transfer functions that decay quickly the delay margin is closely related to the phase margin but for systems where the amplitude ratio of the loop transfer function has several peaks at high frequencies the delay margin is a much more relevant measure.

Example 29 *(Conditional Stability).* Consider a feedback system with the loop transfer function

$$L(s) = \frac{3(s+1)^2}{s(s+6)^2} \tag{7.6}$$

The Nyquist plot of the loop transfer function is shown in Figure 7.8. Notice that the Nyquist curve intersects the negative real axis twice. The first intersection occurs at $s = -$ for $\omega =$ and the second at $s = -$ for $\omega =$. The

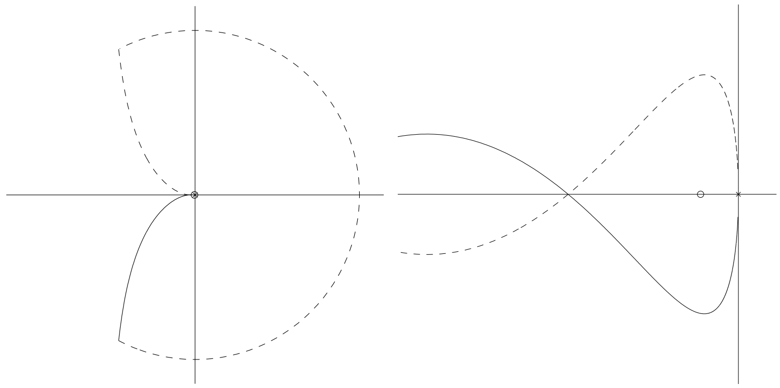

Figure 7.8: Nyquist curve for the loop transfer function $L(s) = \frac{3(s+1)^2}{s(s+6)^2}$. The plot on the right is an enlargement of the area around the origin of the plot on the left.

intuitive argument based on signal tracing around the loop is less intuitive in this case, because injection of a sinusoid with frequency xx and amplitude 1 at A will in steady state give an oscillation at B with amplitude XXX. It follows from Nyquist's stability criterion that the system is stable because the critical point is to the right of the Nyquist curve when it is traversed for increasing frequencies. It was actually systems of this type which motivated much of the research that led Nyquist to develop his stability criterion.

7.5 Second Order Systems

7.6 Further Reading