Design Methods for Control Systems

Okko H. Bosgra
Delft University of Technology
Delft, The Netherlands

Huibert Kwakernaak
University of Twente
Enschede, The Netherlands

Gjerrit Meinsma
University of Twente
Enschede, The Netherlands

Notes for a course of the
Dutch Institute of Systems and Control
Winter term 2002–2003

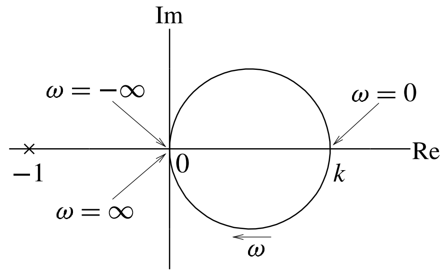

Figure 1.12: Nyquist plot of the loop gain transfer function $L(s) = k/(1 + s\theta)$

This in turn may be represented as

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \mathbf{c} \tag{1.59}$$

and solved for $\mathbf{x}$ and $\mathbf{y}$. If the polynomials D and N are coprime[8] then the square matrix $\begin{bmatrix} A & B \end{bmatrix}$ is nonsingular. □

1.3.5 Nyquist criterion

In classical control theory closed-loop stability is often studied with the help of the *Nyquist stability criterion,* which is a well-known graphical test. Consider the simple MIMO feedback loop of Fig. 1.11. The block marked "L" is the series connection of the compensator C and the plant P. The transfer matrix $L = PK$ is called the *loop gain matrix* — or *loop gain,* for short — of the feedback loop.

For a SISO system, L is a scalar function. Define the *Nyquist plot*[9] of the scalar loop gain L as the curve traced in the complex plane by

$$L(\mathrm{j}\omega), \qquad \omega \in \mathbb{R}. \tag{1.60}$$

Because for finite-dimensional systems L is a rational function with real coefficients, the Nyquist plot is symmetric with respect to the real axis. Associated with increasing ω we may define a positive direction along the locus. If L is *proper*[10] and has no poles on the imaginary axis then the locus is a closed curve. By way of example, Fig. 1.12 shows the Nyquist plot of the loop gain transfer function

$$L(s) = \frac{k}{1 + s\theta}, \tag{1.61}$$

with k and θ positive constants. This is the loop gain of the cruise control system of Example 1.2.5 with $k = g\theta/T$.

We first state the best known version of the Nyquist criterion.

Summary 1.3.10 (Nyquist stability criterion for SISO open-loop stable systems). Assume that in the feedback configuration of Fig. 1.11 the SISO system L is open-loop stable. Then the closed-loop system is stable if and only if the Nyquist plot of L does not encircle the point -1. □

[8] That is, they have no nontrivial common factors.

[9] The Nyquist plot is discussed at more length in § 2.4.3.

[10] A rational matrix function L is *proper* if $\lim_{|s|\to\infty} L(s)$ exists. For a rational function L this means that the degree of its numerator is not greater than that of its denominator.

It follows immediately from the Nyquist criterion and Fig. 1.12 that if $L(s) = k/(1+s\theta)$ and the block "L" is stable then the closed-loop system is stable for all positive k and θ.

Exercise 1.3.11 (Nyquist plot). Verify the Nyquist plot of Fig. 1.12. □

Exercise 1.3.12 (Stability of compensated feedback system). Consider a SISO single-degree-of-freedom system as in Fig. 1.8(a) or (b), and define the loop gain $L = PC$. Prove that if both the compensator and the plant are stable and L satisfies the Nyquist criterion then the feedback system is stable. □

The result of Summary 1.3.10 is a special case of the *generalized Nyquist criterion*. The generalized Nyquist principle applies to a MIMO unit feedback system of the form of Fig. 1.11, and may be phrased as follows:

Summary 1.3.13 (Generalized Nyquist criterion). Suppose that the loop gain transfer function L of the MIMO feedback system of Fig. 1.11 is proper such that $I + L(\mathrm{j}\infty)$ is nonsingular (this guarantees the feedback system to be well-defined) and has no poles on the imaginary axis. Assume also that the Nyquist plot of $\det(I + L)$ does not pass through the origin. Then

the number of unstable closed-loop poles
=
the number of times the Nyquist plot of $\det(I + L)$ encircles the origin clockwise[11]
+
the number of unstable open-loop poles.

It follows that the closed-loop system is stable if and only if the number of encirclements of $\det(I + L)$ equals the negative of the number of unstable open-loop poles. □

Similarly, the "unstable open-loop poles" are the right-half plane eigenvalues of the system matrix of the state space representation of the open-loop system. This includes any uncontrollable or unobservable eigenvalues. The "unstable closed-loop poles" similarly are the right-half plane eigenvalues of the system matrix of the closed-loop system.

In particular, it follows from the generalized Nyquist criterion that if the open-loop system is stable then the closed-loop system is stable if and only if the number of encirclements is zero (i.e., the Nyquist plot of $\det(I + L)$ does *not* encircle the origin).

For SISO systems the loop gain L is scalar, so that the number of times the Nyquist plot of $\det(I + L) = 1 + L$ encircles the origin equals the number of times the Nyquist plot of L encircles the point -1.

The condition that $\det(I + L)$ has no poles on the imaginary axis and does not pass through the origin may be relaxed, at the expense of making the analysis more complicated (see for instance Dorf (1992)).

The proof of the Nyquist criterion is given in § 1.10. More about Nyquist plots may be found in § 2.4.3.

1.3.6 Existence of a stable stabilizing compensator

A compensator that stabilizes the closed-loop system but by itself is unstable is difficult to handle in start-up, open-loop, input saturating or testing situations. There are unstable plants for which a stable stabilizing controller does not exist. The following result was formulated and proved by Youla, Bongiorno, and Lu (1974); see also Anderson and Jury (1976) and Blondel (1994).

[11] This means the number of clockwise encirclements minus the number of anticlockwise encirclements. I.e., this number may be negative.

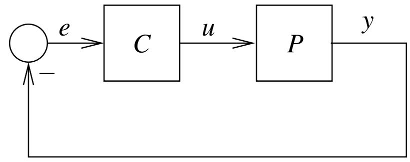

Figure 1.13: Feedback system configuration

Summary 1.3.14 (Existence of stable stabilizing controller). Consider the unit feedback system of Fig. 1.11(a) with plant P and compensator C.

The plant possesses the *parity interlacing property* if it has an even number of poles (counted according to multiplicity) between each pair of zeros on the positive real axis (including zeros at infinity.)

There exists a stable compensator C that makes the closed-loop stable if and only if the plant P has the parity interlacing property. □

If the denominator of the plant transfer function P has degree n and its numerator degree m then the plant has n poles and m (finite) zeros. If $m < n$ then the plant is said to have $n - m$ zeros at infinity.

Exercise 1.3.15 (Parity interlacing property). Check that the plant

$$P(s) = \frac{s}{(s-1)^2} \tag{1.62}$$

possesses the parity interlacing property while

$$P(s) = \frac{(s-1)(s-3)}{s(s-2)} \tag{1.63}$$

does not. Find a stabilizing compensator for each of these two plants (which for the first plant is itself stable.) □

1.4 Stability robustness

1.4.1 Introduction

In this section we consider SISO feedback systems with the configuration of Fig. 1.13. We discuss their *stability robustness,* that is, the property that the closed-loop system remains stable under changes of the plant and the compensator. This discussion focusses on the *loop gain* $L = PC$, with P the plant transfer function, and C the compensator transfer function. For simplicity we assume that the system is *open-loop stable,* that is, both P and C represent the transfer function of a stable system.

We also assume the existence of a *nominal* feedback loop with loop gain L_0, which is the loop gain that is supposed to be valid under nominal circumstances.

1.4.2 Stability margins

The closed-loop system of Fig. 1.13 remains stable under perturbations of the loop gain L as long as the Nyquist plot of the perturbed loop gain does not encircle the point -1. Intuitively, this may be accomplished by "keeping the Nyquist plot of the nominal feedback system away from the point -1."

The classic *gain margin* and *phase margin* are well-known indicators for how closely the Nyquist plot approaches the point -1.

Gain margin The gain margin is the smallest positive number k_m by which the Nyquist plot must be multiplied so that it passes through the point -1. We have

$$k_m = \frac{1}{|L(\mathrm{j}\omega_r)|}, \tag{1.64}$$

where ω_r is the angular frequency for which the Nyquist plot intersects the negative real axis furthest from the origin (see Fig. 1.14).

Phase margin The phase margin is the extra phase ϕ_m that must be added to make the Nyquist plot pass through the point -1. The phase margin ϕ_m is the angle between the negative real axis and $L(\mathrm{j}\omega_m)$, where ω_m is the angular frequency where the Nyquist plot intersects the unit circle closest to the point -1 (see again Fig. 1.14).

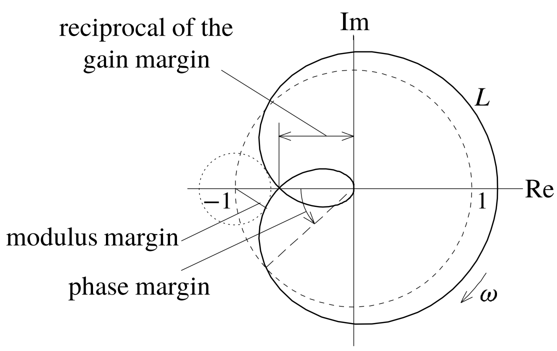

Figure 1.14: Robustness margins

In classical feedback system design, robustness is often specified by establishing minimum values for the gain and phase margin. Practical requirements are $k_m > 2$ for the gain margin and $30^\circ < \phi_m < 60^\circ$ for the phase margin.

The gain and phase margin do not necessarily adequately characterize the robustness. Figure 1.15 shows an example of a Nyquist plot with excellent gain and phase margins but where a relatively small *joint* perturbation of gain and phase suffices to destabilize the system. For this reason Landau, Rolland, Cyrot, and Voda (1993) introduced two more margins.

Modulus margin [12] The modulus margin s_m is the radius of the smallest circle with center -1 that is tangent to the Nyquist plot. Figure 1.14 illustrates this. The modulus margin very directly expresses how far the Nyquist plot stays away from -1.

Delay margin [13] The delay margin τ_m is the smallest extra delay that may be introduced in the loop that destabilizes the system. The delay margin is linked to the phase margin ϕ_m by the relation

$$\tau_m = \min_{\omega_*} \frac{\phi_*}{\omega_*}. \tag{1.65}$$

[12] French: *marge de module.*
[13] French: *marge de retard.*

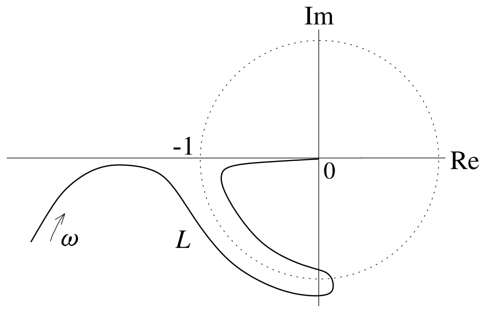

Figure 1.15: This Nyquist plot has good gain and phase margins but a small simultaneous perturbation of gain and phase destabilizes the system

Here ω_* ranges over all nonnegative frequencies at which the Nyquist plot intersects the unit circle, and ϕ_* denotes the corresponding phase $\phi_* = \arg L(\mathrm{j}\omega_*)$. In particular $\tau_m \leq \frac{\phi_m}{\omega_m}$.

A practical specification for the modulus margin is $s_m > 0.5$. The delay margin should be at least of the order of $\frac{1}{2B}$, where B is the bandwidth (in terms of angular frequency) of the closed-loop system.

Adequate margins of these types are not only needed for robustness, but also to achieve a satisfactory time response of the closed-loop system. If the margins are small, the Nyquist plot approaches the point -1 closely. This means that the stability boundary is approached closely, manifesting itself by closed-loop poles that are very near to the imaginary axis. These closed-loop poles may cause an oscillatory response (called "ringing" if the resonance frequency is high and the damping small.)

Exercise 1.4.1 (Relation between robustness margins). Prove that the gain margin k_m and the phase margin ϕ_m are related to the modulus margin s_m by the inequalities

$$k_m \geq \frac{1}{1 - s_m}, \qquad \phi_m \geq 2 \arcsin \frac{s_m}{2}. \tag{1.66}$$

This means that if $s_m \geq \frac{1}{2}$ then $k_m \geq 2$ and $\phi_m \geq 2 \arcsin \frac{1}{4} \approx 28.96^\circ$ (Landau, Rolland, Cyrot, and Voda 1993). The converse is not true in general. □

1.4.3 Robustness for loop gain perturbations

The robustness specifications discussed so far are all rather qualitative. They break down when the system is not open-loop stable, and, even more spectacularly, for MIMO systems. We introduce a more refined measure of stability robustness by considering the effect of plant perturbations on the Nyquist plot more in detail. For the time being the assumptions that the feedback system is SISO and open-loop stable are upheld. Both are relaxed later.

Naturally, we suppose the nominal feedback system to be well-designed so that it is closed-loop stable. We investigate whether the feedback system *remains* stable when the loop gain is perturbed from the nominal loop gain L_0 to the actual loop gain L.

By the Nyquist criterion, the Nyquist plot of the nominal loop gain L_0 does not encircle the point -1, as shown in Fig. 1.16. The actual closed-loop system is stable if also the Nyquist plot of the actual loop gain L does not encircle -1.

It is easy to see by inspection of Fig. 1.16 that the Nyquist plot of L definitely does not encircle the point -1 if for all $\omega \in \mathbb{R}$ the distance $|L(\mathrm{j}\omega) - L_0(\mathrm{j}\omega)|$ between any point $L(\mathrm{j}\omega)$ and the

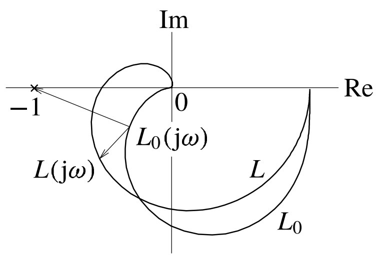

Figure 1.16: Nominal and perturbed Nyquist plots

corresponding point $L_0(j\omega)$ is *less* than the distance $|L_0(j\omega) + 1|$ of the point $L_0(j\omega)$ and the point -1, that is, if

$$|L(j\omega) - L_0(j\omega)| < |L_0(j\omega) + 1| \quad \text{for all } \omega \in \mathbb{R}. \tag{1.67}$$

This is equivalent to

$$\frac{|L(j\omega) - L_0(j\omega)|}{|L_0(j\omega)|} \cdot \frac{|L_0(j\omega)|}{|L_0(j\omega) + 1|} < 1 \quad \text{for all } \omega \in \mathbb{R}. \tag{1.68}$$

Define the *complementary sensitivity function* T_0 of the nominal closed-loop system as

$$T_0 = \frac{L_0}{1 + L_0}. \tag{1.69}$$

T_0 bears its name because its complement

$$1 - T_0 = \frac{1}{1 + L_0} = S_0 \tag{1.70}$$

is the *sensitivity function*. The sensitivity function plays an important role in assessing the effect of disturbances on the feedback system, and is discussed in Section 1.5.

Given T_0, it follows from (1.68) that if

$$\frac{|L(j\omega) - L_0(j\omega)|}{|L_0(j\omega)|} \cdot |T_0(j\omega)| < 1 \quad \text{for all } \omega \in \mathbb{R} \tag{1.71}$$

then the perturbed closed-loop system is stable.

The factor $|L(j\omega) - L_0(j\omega)|/|L_0(j\omega)|$ in this expression is the *relative* size of the perturbation of the loop gain L from its nominal value L_0. The relation (1.71) shows that the closed-loop system is guaranteed to be stable as long as the relative perturbations satisfy

$$\frac{|L(j\omega) - L_0(j\omega)|}{|L_0(j\omega)|} < \frac{1}{|T_0(j\omega)|} \quad \text{for all } \omega \in \mathbb{R}. \tag{1.72}$$

The larger the magnitude of the complementary sensitivity function is, the smaller is the allowable perturbation.

This result is discussed more extensively in Section 5.6, where also its MIMO version is described. It originates from Doyle (1979). The stability robustness condition has been obtained under the assumption that the open-loop system is stable. In fact, it also holds for open-loop unstable systems, *provided* the number of right-half plane poles remains invariant under perturbation.

1.10.1 Closed-loop characteristic polynomial

We first prove (1.44) in Subsection 1.3.3.

Proof 1.10.1 (Closed-loop characteristic polynomial). Let

$$\dot{x} = Ax + Bu, \quad y = Cx + Du, \tag{1.156}$$

be a state realization of the block L in the closed-loop system of Fig. 1.11. It follows that $L(s) = C(sI - A)^{-1} + D$. From $u = -y$ we obtain with the output equation that $u = -Cx - Du$, so that $u = -(I + D)^{-1}Cx$. Since by assumption $I + D = I + L(\mathrm{j}\infty)$ is nonsingular the closed-loop system is well-defined. Substitution of u into the state differential equation shows that the closed-loop system is described by the state differential equation

$$\dot{x} = [A - B(I + D)^{-1}C]x. \tag{1.157}$$

The characteristic polynomial χ_{cl} of the closed-loop system hence is given by

$$\begin{aligned} \chi_{\mathrm{cl}}(s) &= \det[sI - A + B(I + D)^{-1}C] \\ &= \det(sI - A) \cdot \det[I + (sI - A)^{-1}B(I + D)^{-1}C]. \end{aligned} \tag{1.158}$$

Using the well-known determinant equality $\det(I + MN) = \det(I + NM)$ it follows that

$$\begin{aligned} \chi_{\mathrm{cl}}(s) &= \det(sI - A) \cdot \det[I + (I + D)^{-1}C(sI - A)^{-1}B] \\ &= \det(sI - A) \cdot \det[(I + D)^{-1}] \cdot \det[I + D + C(sI - A)^{-1}B] \\ &= \det(sI - A) \cdot \det[(I + D)^{-1}] \cdot \det[I + L(s)]. \end{aligned} \tag{1.159}$$

Denoting the open-loop characteristic polynomial as $\det(sI - A) = \chi(s)$ we thus have

$$\frac{\chi_{\mathrm{cl}}(s)}{\chi(s)} = \frac{\det[I + L(s)]}{\det[I + L(\mathrm{j}\infty)]}. \tag{1.160}$$

■

1.10.2 The Nyquist criterion

The proof of the generalized Nyquist criterion of Summary 1.3.13 in Subsection 1.3.5 relies on the *principle of the argument* of complex function theory[24].

Summary 1.10.2. Principle of the argument Let R be a rational function, and $\mathcal{C}$ a closed contour in the complex plane as in Fig. 1.38. As the complex number s traverses the contour $\mathcal{C}$ in clockwise direction, its image $R(s)$ under R traverses a closed contour that is denoted as $R(\mathcal{C})$, also shown in Fig. 1.38. Then as s traverses the contour $\mathcal{C}$ exactly once in clockwise direction,

(the number of times $R(s)$ encircles the origin in clockwise direction as s traverses $\mathcal{C}$)
=
(the number of zeros of R inside $\mathcal{C}$) $-$ (the number of poles of R inside $\mathcal{C}$).

We prove the generalized Nyquist criterion of Summary 1.3.13.

Proof of the generalized Nyquist criterion. We apply the principle of the argument to (1.160), where we choose the contour $\mathcal{C}$ to be the so-called *Nyquist contour* or *D-contour* indicated in Fig. 1.39. The radius ρ of the semicircle is chosen so large that the contour encloses all the right-half plane roots of both χ_{cl} and χ_{ol}. Then by the principle of the argument the number of times that the image of $\det(I + L)$ encircles the origin equals the number of right-half plane roots of χ_{cl} (i.e., the number of unstable closed-loop poles) minus the number of right-half plane roots of χ_{ol} (i.e., the number of unstable open-loop poles). The Nyquist criterion follows by letting the radius ρ of the semicircle approach ∞. Note that as ρ approaches ∞ the image of the semicircle under $\det(I + L)$ shrinks to the single point $\det(I + L(\mathrm{j}\infty))$. ■

[24] See Henrici (1974). The generalized form in which we state the principle may be found in Postlethwaite and MacFarlane (1979).

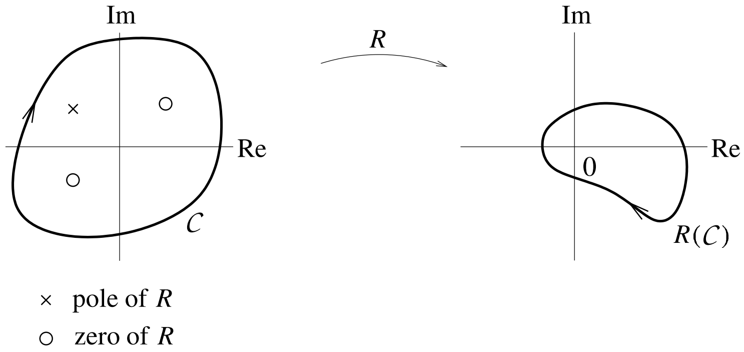

Figure 1.38: Principle of the argument. Left: a closed contour $\mathcal{C}$ in the complex plane. Right: the image $R(\mathcal{C})$ of $\mathcal{C}$ under a rational function R.

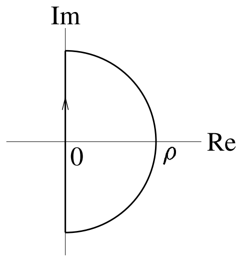

Figure 1.39: Nyquist contour

1.10.3 Bode's sensitivity integral

The proof of Bode's sensitivity integral is postponed until the next subsection. Accepting it as true we use it to derive the inequality (1.117) of Subsection 1.6.3.

Proof 1.10.3 (Lower bound for peak value of sensitivity). If the open-loop system is stable then we have according to Bode's sensitivity integral

$$\int_0^\infty \log|S(\mathrm{j}\omega)|\, d\omega = 0. \tag{1.161}$$

From the assumption that $|S(\mathrm{j}\omega)| \leq \alpha < 1$ for $0 \leq \omega \leq \omega_L$ it follows that if $0 < \omega_L < \omega_H < \infty$ then

$$\begin{aligned} 0 &= \int_0^\infty \log|S(\mathrm{j}\omega)|\, d\omega \\ &= \int_0^{\omega_L} \log|S(\mathrm{j}\omega)|\, d\omega + \int_{\omega_L}^{\omega_H} \log|S(\mathrm{j}\omega)|\, d\omega + \int_{\omega_H}^\infty \log|S(\mathrm{j}\omega)|\, d\omega \\ &\leq \omega_L \log\alpha + (\omega_H - \omega_L) \sup_{\omega_L\leq\omega\leq\omega_H} \log|S(\mathrm{j}\omega)| + \int_{\omega_H}^\infty \log|S(\mathrm{j}\omega)|\, d\omega. \end{aligned} \tag{1.162}$$

As a result,

$$(\omega_H - \omega_L) \sup_{\omega_L\leq\omega\leq\omega_H} \log|S(\mathrm{j}\omega)| \geq \omega_L \log\frac{1}{\alpha} - \int_{\omega_H}^\infty \log|S(\mathrm{j}\omega)|\, d\omega. \tag{1.163}$$