Input-Output Stability

A.R. Teel*, T.T. Georgiou*, L. Praly† and E. Sontag‡

1 Introduction

A common task for an engineer is to design a system which reacts to stimuli in some specific and desirable way. One way to characterize appropriate behavior is through the formalism of input-output stability. In this setting a notion of well-behaved input and output signals is made precise and the question is posed: do well-behaved stimuli (inputs) produce well-behaved responses (outputs)?

General input-output stability analysis has its roots in the development of the electronic feedback amplifier of H.S. Black in 1927 and the subsequent development of classical feedback design tools for linear systems by H. Nyquist and H.W. Bode in the 30's and 40's, all at Bell Telephone Laboratories. These latter tools focused on determining input-output stability of linear feedback systems from the characteristics of the feedback components. Generalizations to nonlinear systems were made by several researchers in the late 50's and early 60's. The most notable contributions were those of G. Zames, then at M.I.T., and I.W. Sandberg at Bell Telephone Laboratories. Indeed, much of this chapter is based on the foundational ideas found in [Sandberg, 1964] and [Zames, 1966], with additional insights drawn from [Safonov, 1980]. A thorough understanding of nonlinear systems from an input-output point of view is still an area of ongoing and intensive research.

The strength of input-output stability theory is that it provides a method for anticipating the qualitative behavior of a feedback system with only rough

*University of Minnesota, Department of Electrical Engineering.

†École Des Mines de Paris, Centre Automatique et Systèmes.

‡Rutgers University, Department of Mathematics.

Submitted to The Control Handbook, *CRC Press*

information about the feedback components. This, in turn, leads to notions of robustness of feedback stability and motivates many of the recent developments in modern control theory.

2 Systems and Stability

Throughout our discussion of input-output stability, a **signal** is a "reasonable" (e.g., piecewise continuous) function defined on a finite or semi-infinite time interval, i.e. an interval of the form $[0, T)$ where T is either a strictly positive real number or infinity. In general, a signal is vector-valued; its components typically represent actuator and sensor values. A **dynamical system** is an object which produces an output signal for each input signal.

To discuss stability of dynamical systems, we introduce the concept of a **norm function**, denoted $||\cdot||$, which captures the "size" of signals defined on the semi-infinite time interval. The significant properties of a norm function are that 1) the norm of a signal is zero if the signal is identically zero, and is a strictly positive number otherwise, 2) scaling a signal results in a corresponding scaling of the norm, and 3) the triangle inequality holds, i.e. $||u_1 + u_2|| \leq ||u_1|| + ||u_2||$. Examples of norm functions are the p**-norms**. For any positive real number $p \geq 1$, the p-norm is defined by

$$||u||_p := \left(\int_0^\infty |u(t)|^p \right)^{\frac{1}{p}} \tag{1}$$

where $|\cdot|$ represents the standard Euclidean norm. For $p = \infty$, we define

$$||u||_\infty := \sup_{t \geq 0} |u(t)| \ . \tag{2}$$

The ∞-norm is useful when amplitude constraints are imposed on a problem, while the 2-norm is of more interest in the context of energy constraints. The norm of a signal may very well be infinite. We will typically be interested in measuring signals which may only be defined on finite time intervals or measuring truncated versions of signals. To that end, given a signal u and a strictly positive real number τ, we use u_τ to denote the **truncated signal** which is generated by appending zeros to u to extend it onto the semi-infinite interval if necessary and then truncating, i.e. u_τ is equal to the (extended) signal on the interval $[0, \tau]$ and is equal to zero on the interval (τ, ∞).

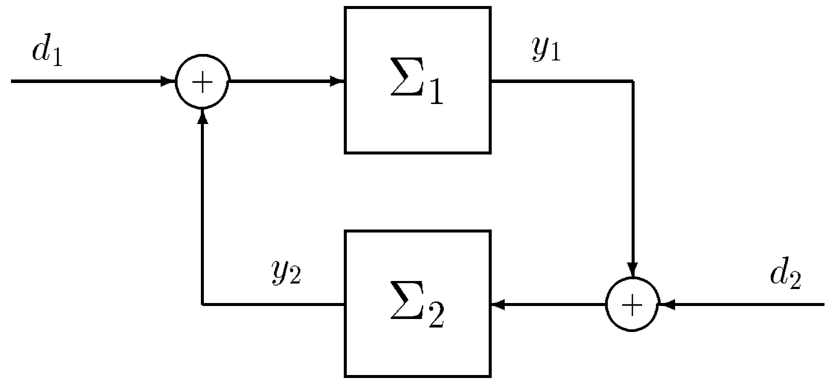

Figure 1: Standard feedback configuration.

Informally, a system is **stable** in the input-output sense if small input signals produce correspondingly small output signals. To make this concept precise, we need a way to quantify the dependence of the norm of the output on the norm of the input applied to the system. To that end, we define a **gain function** as a function from the nonnegative real numbers to the nonnegative real numbers which is continuous, nondecreasing, and zero when its argument is zero. For notational convenience we will say that the "value" of a gain function at ∞ is ∞. A dynamical system is stable (with respect to the norm $||\cdot||$) if there is a gain function γ which gives a bound on the norm of truncated output signals as a function of the norm of truncated input signals, i.e.

$$||y_\tau|| \leq \gamma(||u_\tau||) \qquad \text{for all } \tau \ . \tag{3}$$

In the very special case when the gain function is linear, i.e. there is at most an amplification by a constant factor, the dynamical system is **finite gain stable**. The notions of finite gain stability and closely related variants are central to much of classical input-output stability theory, but in recent years much progress has been made as well in understanding the role of more general (nonlinear) gains in system analysis.

The focus of this chapter will be on the stability analysis of interconnected dynamical systems as described in figure 1. The composite system in figure 1 will be called a **well-defined interconnection** if it is a dynamical system with $\binom{d_1}{d_2}$ as input and $\binom{y_1}{y_2}$ as output. The dynamical systems which make up a well-defined interconnection will be called its **feedback components**. For stability of well-defined interconnections, it is not necessary for either of the feedback components to be stable nor is it sufficient for both of the feedback components to be stable. On the other hand, necessary and sufficient conditions for stability of a well-defined interconnection can be expressed in terms of the set of all possible input-output pairs for each feedback component. To be explicit in this regard, we make some definitions. For a given dynamical system Σ with input signals u and output signals y, the set of its ordered input-output pairs (u, y) is referred to as the **graph** of the dynamical system and is denoted $\mathcal{G}_\Sigma$. When the input and output are exchanged in the ordered pair, i.e. (y, u), the set is referred to as the **inverse graph** of the system and is denoted $\mathcal{G}^I_\Sigma$. Note that, for the system in figure 1, the inverse graph of Σ_2 and the graph of Σ_1 lie in the same Cartesian product space which we will call the **ambient space.** We will use as norm on the ambient space the sum of the norms of the coordinates.

The basic observation regarding input-output stability for a well-defined interconnection says, in informal terms, that if a signal in the inverse graph of Σ_2 is near any signal in the graph of Σ_1 then it must be small. To formalize this notion, we need the concept of the **distance** to the graph of Σ_1 from signals x in the ambient space. This (truncated) distance is defined by

$$d_\tau(x, \mathcal{G}_{\Sigma_1}) := \inf_{z \in \mathcal{G}_{\Sigma_1}} ||(x - z)_\tau|| \ . \tag{4}$$

Graph separation theorem: *A well-defined interconnection is stable if and only if there exists a gain function γ which gives a bound on the norm of truncated signals in the inverse graph of Σ_2 as a function of the truncated distance from the signals to the graph of Σ_1, i.e.*

$$x \in \mathcal{G}^I_{\Sigma_2} \qquad \Longrightarrow \qquad ||x_\tau|| \leq \gamma \Big(\ d_\tau \left(x, \mathcal{G}_{\Sigma_1}\right) \ \Big) \qquad \text{for all } \tau \ . \tag{5}$$

In the special case where γ is a linear function, the well-defined interconnection is finite gain stable.

The idea behind this observation can be understood by considering the signals that arise in the closed loop which belong to the inverse graph of Σ_2, i.e. the signals $(y_2, y_1 + d_2)$. (Stability with these signals taken as output is equivalent to stability with the original outputs.) Notice that, for the system in figure 1, signals in the graph of Σ_1 have the form (y_2+d_1, y_1). Consequently, signals $x \in \mathcal{G}^I_{\Sigma_2}$ and $z \in \mathcal{G}_{\Sigma_1}$ which satisfy the feedback equations also satisfy

$$(x - z)_\tau = (d_1, -d_2)_\tau \tag{6}$$

and

$$||(x - z)_\tau|| = \|(d_1, d_2)_\tau\| \tag{7}$$

for truncations within the interval of definition. If there are signals x in the inverse graph of Σ_2 with large truncated norm but small truncated distance to the graph of Σ_1, i.e. there exists some $z \in \mathcal{G}_{\Sigma_1}$ and $\tau > 0$ such that $||(x-z)_\tau||$ is small, then we can choose (d_1, d_2) to satisfy (6) giving, according to (7), a small input which produces a large output. This contradicts our definition of stability. Conversely, if there is no z which is close to x then only large inputs can produce large x signals and thus the system is stable.

The distance observation presented above is the unifying idea behind the input-output stability criteria that are applied in practice. However, the observation is rarely applied directly because of the difficulties involved in exactly characterizing the graph of a dynamical system and measuring distances. Instead, various simpler conditions have been developed which constrain the graphs of the feedback components in a way that guarantees the graph of Σ_1 and the inverse graph of Σ_2 are sufficiently separated. There are many such sufficient conditions and in the remainder of this chapter we will describe a few of them.

3 Practical conditions and examples

3.1 The classical small gain theorem

One of the most commonly used sufficient conditions for graph separation constrains the graphs of the feedback components by assuming that each feedback component is finite gain stable. Then, the appropriate graphs will be separated if the product of the coefficients of the linear gain functions is sufficiently small. For this reason, the result based on this type of constraint has come to be known as the small gain theorem.

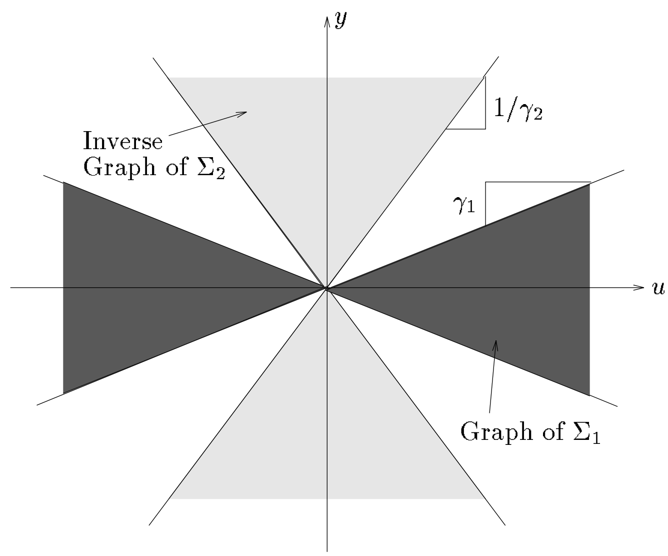

Figure 2: Classical small gain theorem

Small gain theorem: *If each feedback component is finite gain stable and the product of the gains (the coefficients of the linear gain functions) is less than one then the well-defined interconnection is finite gain stable.*

Figure 2 provides the intuition for the result. If we were to draw an analogy between a dynamical system and a static map whose graph is a set of points in the plane, the graph of Σ_1 would be constrained to the darkly shaded conic region by the finite gain stability assumption. Likewise, the inverse graph of Σ_2 would be constrained to the lightly shaded region. The fact that the product of the gains is less than one guarantees the positive aperture between the two regions and, in turn, that the graphs are separated sufficiently.

To apply the small gain theorem, we need a way to verify that the feedback components are finite gain stable (with respect to a particular norm) and determine their gains. In particular, any linear dynamical system that can be represented with a real, rational transfer function $G(s)$ is finite gain stable in any of the p-norms if and only if all of the poles of the transfer function have negative real part. A popular norm to work with is the 2-norm. It is associated with the energy of a signal. For a single-input, single-output (SISO) finite gain stable system modeled by a real, rational transfer function $G(s)$, the smallest possible coefficient for the stability gain function with respect to the 2-norm is given by:

$$\bar{\gamma} := \sup_{\omega} |\mathrm{G}(j\omega)| \ . \tag{8}$$

For multi-input, multi-output systems, the magnitude in (8) is replaced by the maximum singular value. In either case, this can be established using **Parseval's theorem**. For SISO systems, the quantity in (8) can be obtained from a quick examination of the Bode plot or Nyquist plot for the transfer function. If the Nyquist plot of a stable SISO transfer function lies inside a circle of radius $\bar{\gamma}$ centered at the origin then the coefficient of the 2-norm gain function for the system is less than or equal to $\bar{\gamma}$.

More generally, consider a dynamical system that can be represented by a finite dimensional ordinary differential equation with zero initial state :

$$\begin{array}{rcll} \dot{x} & = & f(x,u) \quad , & \quad x(0)=0 \\ y & = & h(x,u) \quad . & \end{array} \tag{9}$$

Suppose that f has globally bounded partial derivatives and that there exist positive real numbers ℓ_1 and ℓ_2 such that

$$|h(x,u)| \leq \ell_1|x| + \ell_2|u| \ . \tag{10}$$

Under these conditions, if the trajectories of the unforced system with nonzero initial conditions,

$$\dot{x} = f(x,0) \quad , \qquad x(0) = x_\circ \ , \tag{11}$$

satisfy

$$|x(t)| \leq k \exp(-\lambda t)|x_\circ| \tag{12}$$

for some positive real number k and λ and any $x_\circ \in \mathbb{R}^n$ then the system (9) is finite gain stable in any of the p-norms. This can be established using Lyapunov function arguments that apply to the system (11).

Example 3.1 Consider a nonlinear control system modeled by an ordinary differential equation with state $x \in \mathbb{R}^n$, input $v \in \mathbb{R}^m$ and disturbance $d_1 \in \mathbb{R}^m$:

$$\dot{x} = f(x, v + d_1) \ . \tag{13}$$

Suppose that f has globally bounded partial derivatives and that a control $v = \alpha(x)$ can be found, also with a globally bounded partial derivative, so that the trajectories of the system

$$\dot{x} = f(x, \alpha(x)) \quad , \qquad x(0) = x_\circ \tag{14}$$

satisfy the bound

$$|x(t)| \le k \exp(-\lambda t)|x_\circ| \tag{15}$$

for some positive real numbers k and λ and for all $x_\circ \in \mathbb{R}^n$. As mentioned above, for any function h satisfying the type of bound in (10), this implies that the system

$$\begin{aligned} \dot{x} &= f(x, \alpha(x) + d_1) \quad , \qquad x(0) = 0 \\ y &= h(x, d_1) \end{aligned} \tag{16}$$

has finite 2-norm gain from input d_1 to output y. We consider the output

$$y := \dot{\alpha} = \frac{\partial \alpha}{\partial x} f(x, \alpha(x) + d_1) \tag{17}$$

which satisfies the type of bound in (10) since α and f both have globally bounded partial derivatives.

We will show, using the small gain theorem, that disturbances d_1 with finite 2-norm continue to produce outputs y with finite 2-norm even when the actual input v to the process is generated by the following fast dynamic version of the commanded input $\alpha(x)$:

$$\begin{aligned} \epsilon \dot{z} &= Az + B(\alpha(x) + d_2) \quad , \qquad z(0) = -A^{-1}B\alpha(x(0)) \\ v &= Cz \ . \end{aligned} \tag{18}$$

Here, ϵ is a small positive parameter, the eigenvalues of A all have strictly negative real part (thus A is invertible) and $-CA^{-1}B = I$. This system may represent unmodeled actuator dynamics.

To see the stability result, we will consider the composite system in the coordinates x and $\zeta = z + A^{-1}B\alpha(x)$. We have, using the notation from figure 1 :

$$\Sigma_1 : \qquad \begin{aligned} \dot{x} &= f(x, \alpha(x) + u_1) \quad , \qquad x(0) = 0 \\ y_1 &= \dot{\alpha}(x) \end{aligned} \tag{19}$$

and

$$\Sigma_2 : \qquad \begin{aligned} \dot{\zeta} &= \epsilon^{-1}A\zeta + A^{-1}Bu_2 \quad , \qquad \zeta(0) = 0 \\ y_2 &= C\zeta \end{aligned} \tag{20}$$

with the interconnection conditions

$$u_1 = y_2 + d_1, \qquad u_2 = y_1 + \epsilon^{-1}Ad_2 \ . \tag{21}$$

Of course, if the system is finite gain stable with the inputs d_1 and $\epsilon^{-1}Ad_2$ then it is also finite gain stable with the inputs d_1 and d_2. We have already discussed that the system Σ_1 in (19) has finite 2-norm gain, say γ_1. Now consider the system Σ_2 in (20). It can be represented with the transfer function

$$\begin{aligned} G(s) &= C(sI - \epsilon^{-1}A)^{-1}A^{-1}B \\ &= \epsilon C(\epsilon sI - A)^{-1}A^{-1}B \\ &=: \epsilon \bar{G}(\epsilon s) \ . \end{aligned} \tag{22}$$

Identifying $\bar{G}(s) = C(sI - A)^{-1}A^{-1}B$, we see that if

$$\gamma_2 := \sup_{\omega} \sigma(\bar{G}(j\omega)) \tag{23}$$

then

$$\sup_{\omega} \sigma(G(j\omega)) = \epsilon\gamma_2 \ . \tag{24}$$

We conclude from the small gain theorem that if $\epsilon < \frac{\gamma_1}{\gamma_2}$ then the composite system (19)-(21) with inputs d_1 and d_2 and outputs $y_1 = \dot{\alpha}(x)$ and $y_2 = C\zeta$ is finite gain stable.

3.2 The classical passivity theorem

Another very popular condition which is used to guarantee graph separation is given in the *passivity theorem.* For the most straightforward passivity result, we must have that the number of input channels is equal to the number of output channels for each feedback component. We then identify the relative location of the graphs of the feedback components using a condition involving the integral of the product of the input and the output signals. This operation is known as the **inner product** and is denoted $\langle \cdot, \cdot \rangle$. In particular, for two signals u and y of the same dimension defined on the semi-infinite interval,

$$\langle u, y \rangle := \int_0^\infty u^T(t)y(t)dt \ . \tag{25}$$

Note that $\langle u, y \rangle = \langle y, u \rangle$ and $\langle u, u \rangle = ||u||_2^2$. A dynamical system is **passive** if, for each input-output pair (u, y) and each $\tau > 0$, $\langle u_\tau, y_\tau \rangle \geq 0$. The terminology used here comes from the special case where the input and output are a voltage and a current, respectively, and the energy absorbed by the dynamical system, which is the inner product of the input and output, is nonnegative.

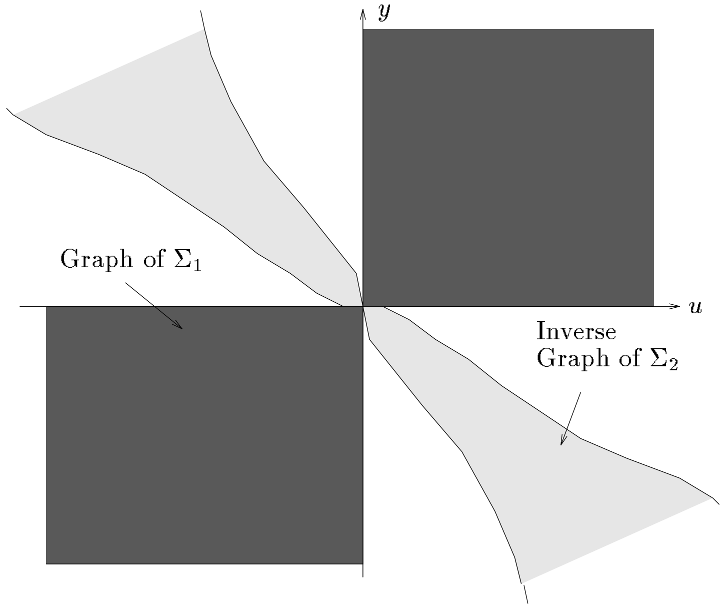

Figure 3: General passivity-based interconnection

Again by analogy to a static map whose graph lies in the plane, passivity of a dynamical system can be thought of as the condition that the graph is constrained to the darkly shaded region in figure 3, i.e. the first and third quadrant of the plane. This graph and the inverse graph of a second system would be separated if, for example, the inverse graph of the second system were constrained to the lightly shaded region in 3, i.e. the second and fourth quadrant but bounded away from the horizontal and vertical axes by an increasing and unbounded distance. But, this is the same as asking that the graph of the second system followed by the scaling '-1', i.e. all pairs $(u, -y)$, be constrained to the first and third quadrant again bounded away from the axes by an increasing and unbounded distance, as in figure 4a. For classical passivity theorems, this region is given a linear boundary as in figure 4b. Notice that, for points $(u_\circ, y_\circ)$ in the plane, if $u_\circ \cdot y_\circ \geq \epsilon(u_\circ^2 + y_\circ^2)$ then $(u_\circ, y_\circ)$ is in the first or third quadrant and $(\epsilon)^{-1}|u_\circ| \geq |y_\circ| \geq \epsilon|u_\circ|$ as in figure 4b. This motivates the following stronger version of passivity. A dynamical system is **input and output strictly passive** if there exists a

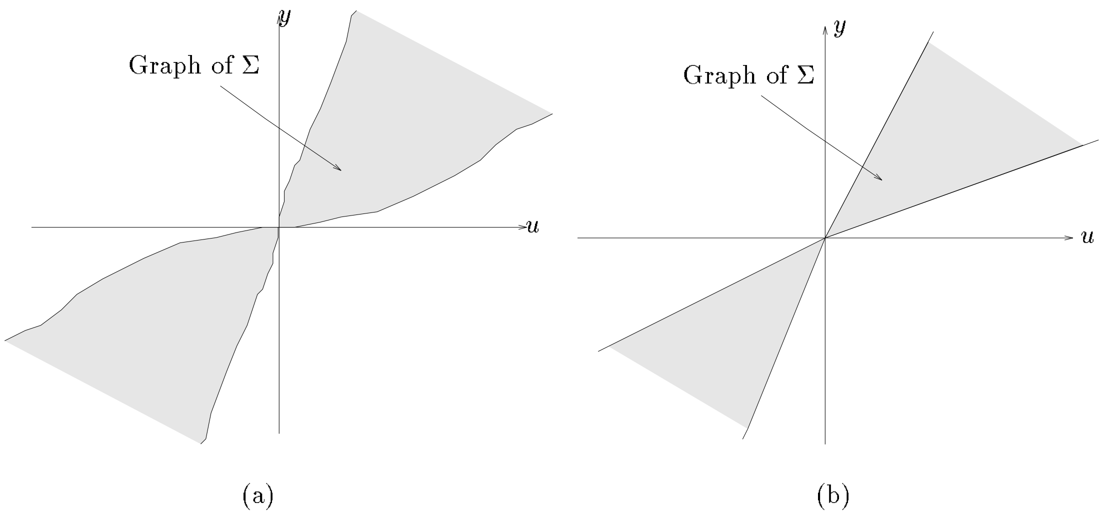

Figure 4: Different notions of input and output strict passivity

strictly positive real number ϵ such that, for each input-output pair (u, y) and each $\tau > 0$, $\langle u_\tau, y_\tau \rangle \geq \epsilon \, (||u_\tau||_2^2 + ||y_\tau||_2^2)$.

There are intermediate versions of passivity which are also useful. These correspond to asking for an increasing and unbounded distance from either the horizontal axis or the vertical axis but not both. For example, a dynamical system is **input strictly passive** if there exists a strictly positive real number ϵ such that, for each input-output pair (u, y) and each $\tau > 0$, $\langle u_\tau, y_\tau \rangle \geq \epsilon ||u_\tau||_2^2$. Similarly, a dynamical system is **output strictly passive** if there exists a strictly positive real number ϵ such that, for each input-output pair (u, y) and each $\tau > 0$, $\langle u_\tau, y_\tau \rangle \geq \epsilon ||y_\tau||_2^2$. It is worth noting that input and output strict passivity is equivalent to input strict passive plus finite gain stability. This can be shown with standard manipulations of the inner product. Also, the reader is warned that all three types of strict passivity mentioned above are frequently simply called "strict passivity" in the literature.

Again by thinking of a graph of a system as a set of points in the plane, output strict passivity is the condition that the graph is constrained to the darkly shaded region in figure 5, i.e. the first and third quadrant with an increasing and unbounded distance from the vertical axis. To complement

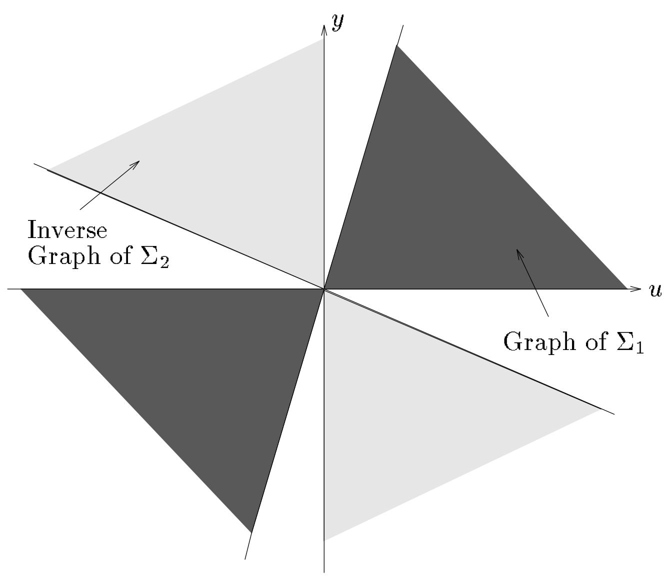

Figure 5: Interconnection of output strictly passive systems

such a graph, consider a second dynamical system which, when followed by the scaling '-1', is also output strictly passive. Such a system has a graph (without the '-1' scaling) which is constrained to the second and fourth quadrant with an increasing and unbounded distance from the vertical axis. In other words, its inverse graph is constrained to the lightly shaded region of figure 5, i.e. to the second and fourth quadrants but with an increasing and unbounded distance from the *horizontal* axis. The conclusions that we can then draw, using the graph separation theorem, are summarized in the following passivity theorem.

Passivity theorem: *If one dynamical system and the other dynamical system followed by the scaling '-1' are*

- *both input strictly passive, OR*
- *both output strictly passive, OR*
- *respectively, passive* and *input and output strictly passive*

then the well-defined interconnection is finite gain stable in the 2-norm.

To apply this theorem, we need a way to verify that the (possibly scaled) feedback components are appropriately passive. For stable SISO systems with real, rational transfer function $G(s)$ it again follows from Parseval's theorem that if

$$\text{Re } G(j\omega) \geq 0$$

for all real values of ω then the system is passive. If the quantity $\text{Re}\, G(j\omega)$ is positive and uniformly bounded away from zero for all real values of ω then the linear system is input and output strictly passive. Similarly, if there exists $\epsilon > 0$ such that, for all real values of ω,

$$\text{Re } G(j\omega - \epsilon) \geq 0 \tag{26}$$

then the linear system is output strictly passive. So, for SISO systems modeled with real, rational transfer functions, passivity and the various forms of strict passivity can again be easily checked by means of a graphical approach such as a Nyquist plot.

More generally, for any dynamical system that can be modeled with a smooth, finite dimensional ordinary differential equation

$$\begin{array}{rcl} \dot{x} & = & f(x) + g(x)u \quad , \qquad x(0) = 0 \\ y & = & h(x) \ , \end{array} \tag{27}$$

if there exists a strictly positive real number ϵ and a nonnegative function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ with $V(0) = 0$ satisfying

$$\frac{\partial V}{\partial x}(x) f(x) \;\leq\; -\epsilon h^T(x) h(x) \tag{28}$$

$$\frac{\partial V}{\partial x}(x) g(x) \;=\; h^T(x) \tag{29}$$

then the system is output strictly passive. With $\epsilon = 0$, the system is passive. Both of these results are established by integrating $\dot{V}$ over the semi-infinite interval.

Example 3.2 (This example is motivated by the work in Berghuis and Nijmeijer, *Systems & Control Letters*, 21:289–295. 1993). Consider a "completely controlled dissipative Euler-Lagrange" system with generalized "forces" F, generalized coordinates q, uniformly positive definite "inertia" matrix $I(q)$, Rayleigh dissipation function $R(\dot{q})$ and, say positive, potential $V(q)$ starting from the position q_d. Namely, let the dynamics of the system be given by the Euler-Lagrange-Rayleigh equations :

$$\dot{\overbrace{\frac{\partial L}{\partial \dot{q}}(q,\dot{q})}} \;=\; \frac{\partial L}{\partial q}(q,\dot{q}) \;+\; F^\top \;-\; \frac{\partial R}{\partial \dot{q}}(\dot{q}) \quad , \qquad q(0) = q_d \ , \ \dot{q}(0) = 0 \tag{30}$$

where L is the Lagrangian :

$$L(q,\dot{q}) \;=\; \tfrac{1}{2}\, \dot{q}^\top I(q)\, \dot{q} \;-\; V(q) \ . \tag{31}$$

Along the solutions of (30), we have :

$$\dot{L} \;=\; \frac{\partial L}{\partial \dot{q}}\ddot{q} + \frac{\partial L}{\partial q}\dot{q} \;=\; \frac{\partial L}{\partial \dot{q}}\ddot{q} + \left[\dot{\overbrace{\frac{\partial L}{\partial \dot{q}}}} - F^\top + \frac{\partial R}{\partial \dot{q}} \right] \dot{q} \tag{32}$$

$$=\; \dot{\overbrace{\left(\frac{\partial L}{\partial \dot{q}}\dot{q} \right)}} - \left[F^\top - \frac{\partial R}{\partial \dot{q}} \right] \dot{q} \;=\; \dot{\overbrace{\left(\dot{q}^\top I(q)\, \dot{q} \right)}} - \left[F^\top - \frac{\partial R}{\partial \dot{q}} \right] \dot{q} \tag{33}$$

$$=\; 2\,\dot{\overbrace{L + V}} - \left[F^\top - \frac{\partial R}{\partial \dot{q}} \right] \dot{q} \;=\; -2\,\dot{V} + \left[F^\top - \frac{\partial R}{\partial \dot{q}} \right] \dot{q} \tag{34}$$

We will suppose that there exists $\epsilon > 0$ such that

$$\frac{\partial R}{\partial \dot{q}}(\dot{q})\,\dot{q} \;\geq\; \epsilon\,|\dot{q}|^2 \;. \tag{35}$$

Now let V_d be a function so that the modified potential:

$$V_m = V + V_d \tag{36}$$

has a global minimum as $q = q_d$ and let the generalized "force" be

$$F \;=\; -\frac{\partial V_d}{\partial q}(q) \;+\; F_m \;. \tag{37}$$

We can see that the system (30) combined with (37), having input F_m and output $\dot{q}$, is output strictly passive by integrating the derivative of the defined Hamiltonian

$$H \;=\; \tfrac{1}{2}\,\dot{q}^{\top} I(q)\,\dot{q} \;+\; V_m(q) \;=\; L \;+\; 2\,V + V_d \;. \tag{38}$$

Indeed the derivative is

$$\dot{H} \;=\; \left[F_m^{\top} - \frac{\partial R}{\partial \dot{q}}\right]\dot{q} \tag{39}$$

and, integrating, we get for each τ :

$$\left\langle \left[F_{m_\tau} - \frac{\partial R}{\partial \dot{q}}(\dot{q}_\tau)^{\top}\right], \dot{q}_\tau \right\rangle \;=\; H(\tau) - H(0)\;. \tag{40}$$

Since $H \geq 0$, $H(0) = 0$ and (35) holds we get

$$\langle F_{m\tau}, \dot{q}_\tau\rangle \;\geq\; \epsilon||\dot{q}_\tau||_2^2 \;. \tag{41}$$

Now we know from the passivity theorem that if F_m is generated from negative dynamic feedback of $\dot{q}$, where the compensator is output strictly passive, then the composite feedback system, with the standard definitions of inputs and outputs, will be finite gain stable using the 2-norm. One such compensator is simply the identity mapping. However, there is interest in designing a linear, output strictly passive compensator which, in addition, has no direct feed-through term. The reason is that if, in fact, there is no disturbance at the input of the compensator we can implement the compensator with measurement only of q and without $\dot{q}$. Indeed, in general

$$G(s)\dot{q} = G(s)s\left(\frac{1}{s}\dot{q}\right) = G(s)s(q - q_d) \tag{42}$$

and the system $G(s)s$ is implementable if $G(s)$ has no direct feed-through terms. To design an output strictly passive linear system without direct feed-through, let A be a matrix having all eigenvalues with strictly negative real part so that, by a well-known result in linear systems theory, there exists a positive definite matrix P satisfying :

$$A^T P + PA = -I\;. \tag{43}$$

Then, for any B matrix of appropriate dimensions, the system modeled by the transfer function

$$G(s) = B^T P(sI - A)^{-1} B \tag{44}$$

is output strictly passive. To see this, consider a state-space realization

$$\begin{aligned} \dot{x} &= Ax + Bu \qquad x(0) = 0 \\ y &= B^T P x \end{aligned} \tag{45}$$

and note that

$$\overset{\cdot}{\overbrace{x^\top P x}} = -x^\top x + 2x^\top P B u \tag{46}$$

$$= -x^\top x + 2y^\top u \; . \tag{47}$$

But, with (45) we have, for some strictly positive real number c,

$$2c\, y^\top y \le x^\top x \; . \tag{48}$$

So, integrating (47) and using that P is positive definite, we get, for all τ,

$$\langle y_\tau, u_\tau \rangle \ge c||y_\tau||_2^2 \; . \tag{49}$$

As a point of interest, one could verify that

$$G(s)s = B^T P A(sI - A)^{-1} B + B^T P B \; . \tag{50}$$

3.3 Simple nonlinear separation theorems

In this section we illustrate how allowing regions with nonlinear boundaries in the small gain and passivity contexts may be useful. First we need a class of functions that will help us describe nonlinear boundaries. A **proper separation function** is a function from the nonnegative real numbers to the nonnegative real numbers which is continuous, zero at zero, strictly increasing and unbounded. The main difference between a gain function and a proper separation function is that the latter is invertible and the inverse is another proper separation function.

3.3.1 Nonlinear passivity

We will briefly discuss a notion of nonlinear input and output strict passivity. To our knowledge, this notion has not been used much in the literature. The notion that we have in mind simply replaces the linear boundaries in the input and output strict passivity definition by nonlinear boundaries as in figure 4a. A dynamical system is **nonlinearly input and output strictly passive** if there exists a proper separation function ρ such that, for each input-output pair (u, y) and each $\tau > 0$, $\langle u_\tau, y_\tau \rangle \geq ||u_\tau||_2 \rho(||u_\tau||_2) + ||y_\tau||_2 \rho(||y_\tau||_2)$. (Note that in the classical definition of strict passivity, $\rho(s) = \epsilon s$ for all $s \geq 0$.)

> **Nonlinear passivity theorem:** *If one dynamical system is passive and the other dynamical system followed by the scaling '-1' is nonlinearly input and output strictly passive then the well-defined interconnection is stable using the 2-norm.*

Example 3.3 Let Σ_1 be a single integrator system:

$$\begin{array}{rcll} \dot{x}_1 & = & u_1 & x_1(0) = 0 \\ y_1 & = & x_1 \ . & \end{array} \tag{51}$$

This system is passive since

$$0 \leq \tfrac{1}{2} x_1(\tau)^2 = \int_0^\tau \frac{d}{dt} \tfrac{1}{2} x(t)^2 dt = \int_0^\tau y_1(t) u_1(t) dt = \langle y_{1_\tau}, u_{1_\tau} \rangle \ . \tag{52}$$

Let the Σ_2 be a system which scales the instantaneous value of the input according to the energy of the input:

$$\begin{array}{rcll} \dot{x}_2 & = & u_2^2 & x_2(0) = 0 \\ y_2 & = & -u_2 \left(\dfrac{1}{1 + |x_2|^{0.25}} \right) \ . & \end{array} \tag{53}$$

This system followed by the scaling '-1' is nonlinearly strictly passive. To see this, first note that

$$x_2(t) = ||u_{2_t}||_2^2 \tag{54}$$

which is a nondecreasing function of t. So,

$$\begin{array}{rcl} \langle -y_{2_\tau}, u_{2_\tau} \rangle & = & \displaystyle\int_0^\tau u_2^2(t) \left(\frac{1}{1 + |x_2(t)|^{0.25}} \right) dt \\ & \geq & \displaystyle\left(\frac{1}{1 + |x_2(\tau)|^{0.25}} \right) \int_0^\tau u_2^2(t) dt \\ & = & \displaystyle\left(\frac{1}{1 + ||u_{2_\tau}||_2^{0.5}} \right) ||u_{2_\tau}||_2^2 \ . \end{array} \tag{55}$$

Now we can define

$$\rho(s) := \frac{0.5 s}{1 + s^{0.5}} \tag{56}$$

which is a proper separation function so that

$$\langle -y_{2_\tau}, u_{2_\tau} \rangle \geq 2\rho(||u_{2_\tau}||_2)||u_{2_\tau}||_2 \ . \tag{57}$$

Finally, note that

$$||y_{2_\tau}||_2^2 = \int_0^\tau u_2^2(t) \frac{1}{\left(1 + x_2^{0.25}(t)\right)^2} dt \leq ||u_{2_\tau}||_2^2 \tag{58}$$

so that

$$\langle -y_{2_\tau}, u_{2_\tau} \rangle \geq \rho(||u_{2_\tau}||_2)||u_{2_\tau}||_2 + \rho(||y_{2_\tau}||_2)||y_{2_\tau}||_2 \ . \tag{59}$$

The conclusion that we can then draw from the nonlinear passivity theorem is that the interconnection of these two systems:

$$\begin{aligned} \dot{x_1} &= -(x_1 + d_2)\left(\frac{1}{1 + |x_2|^{0.25}}\right) + d_1 \\ \dot{x_2} &= (x_1 + d_2)^2 \\ y_1 &= x_1 \\ y_2 &= -(x_1 + d_2)\left(\frac{1}{1 + |x_2|^{0.25}}\right) \end{aligned} \tag{60}$$

is stable when measuring input (d_1, d_2) and output (y_1, y_2) using the 2-norm.

3.3.2 Nonlinear small gain

Just as with passivity, the idea behind the small gain theorem does not require the use of linear boundaries. Consider a well-defined interconnection where each feedback component is stable but not necessarily finite gain stable. Let γ_1 be a stability gain function for Σ_1 and let γ_2 be a stability gain function for Σ_2. Then the graph separation condition will be satisfied if the distance between the curves $(s, \gamma_1(s))$ and $(\gamma_2(r), r)$ grows without bound as in figure 6. This is equivalent to asking whether it is possible to add to the curve $(s, \gamma_1(s))$ in the vertical direction and to the curve $(\gamma_2(r), r)$ in the horizontal direction by an increasing and unbounded amount, i.e. yielding new curves $\left(s, \gamma_1(s) + \rho(s) \right)$ and $\left(\gamma_2(r) + \rho(r), r \right)$ where ρ is a proper separation function, in such a way that the modified first curve is never above the modified second curve. If this is possible, we will say that the composition of the functions γ_1 and γ_2 is a **strict contraction**. To say that a curve $(s, \tilde{\gamma}_1(s))$ is never above a second curve $(\tilde{\gamma}_2(r), r)$ is equivalent to saying that $\tilde{\gamma}_1(\tilde{\gamma}_2(s)) \leq s$ or $\tilde{\gamma}_2(\tilde{\gamma}_1(s)) \leq s$ for all $s \geq 0$. (Equivalently, we

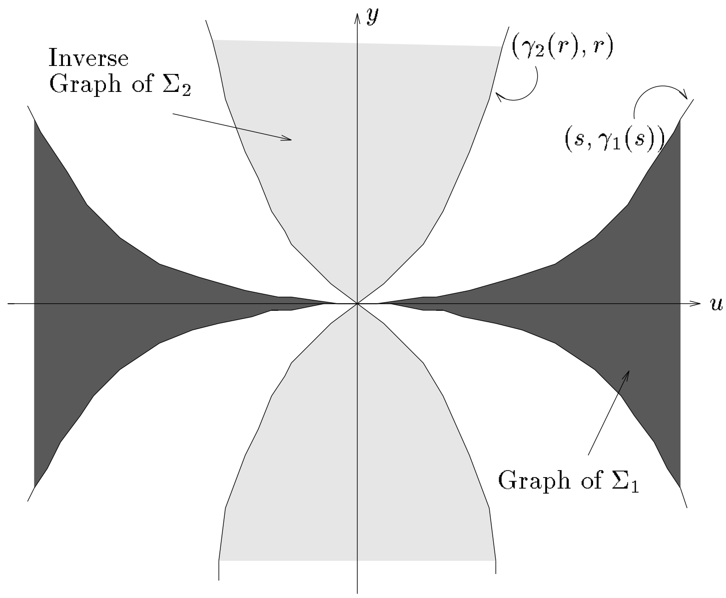

Figure 6: Nonlinear small gain theorem

will write $\tilde{\gamma}_1 \circ \tilde{\gamma}_2 \leq \mathrm{Id}$ or $\tilde{\gamma}_2 \circ \tilde{\gamma}_1 \leq \mathrm{Id}$.) So, requiring that the composition of γ_1 and γ_2 is a strict contraction is equivalent to requiring that there exists a strictly proper separation function ρ so that $(\gamma_1 + \rho) \circ (\gamma_2 + \rho) \leq \mathrm{Id}$, (equivalently $(\gamma_2 + \rho) \circ (\gamma_1 + \rho) \leq \mathrm{Id}$). This condition was made precise in [Mareels and Hill, 1992]. (See also [Jiang, et. al., 1995].) Note that it is not enough to add to just one curve since it is possible for the vertical or horizontal distance to grow without bound while the total distance remains bounded. Finally, note that if the gain functions are linear the condition is the same as the condition that the product of the gains is less than one.

Nonlinear small gain theorem: *If each feedback component is stable (with gain functions γ_1 and γ_2) and the composition of the gains is a strict contraction then the well-defined interconnection is stable.*

To apply the nonlinear small gain theorem we need a way to verify that the feedback components are stable. To date, the most common setting for using the nonlinear small gain theorem is when measuring the input and output using the ∞-norm. For a nonlinear system which can be represented by a smooth, ordinary differential equation :

$$\begin{array}{rcll} \dot{x} & = & f(x,u) \quad , & x(0)=0 \\ y & = & h(x,u) & \end{array} \tag{61}$$

where $h(0,0)=0$, the system is stable (with respect to the ∞-norm) if there exists a positive definite and radially unbounded function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, a proper separation function ψ and a gain function $\tilde{\gamma}$ such that

$$\frac{\partial V}{\partial x} f(x,u) \leq -\psi(|x|) + \tilde{\gamma}(|u|) \ . \tag{62}$$

Since V is positive definite and radially unbounded, there exist additional proper separation functions $\underline{\alpha}$ and $\bar{\alpha}$ such that

$$\underline{\alpha}(|x|) \leq V(x) \leq \bar{\alpha}(|x|) \ . \tag{63}$$

Also, since h is continuous and zero at zero, there exist gain functions ϕ_x and ϕ_u such that

$$|h(x,u)| \leq \phi_x(|x|) + \phi_u(|u|) \ . \tag{64}$$

Given all of these functions, a stability gain function can be computed to be

$$\gamma = \phi_x \circ \underline{\alpha}^{-1} \circ \bar{\alpha} \circ \psi^{-1} \circ \tilde{\gamma} + \phi_u \ . \tag{65}$$

For more details, the reader is directed to [Sontag, 1989].

Example 3.4 Consider the composite system

$$\begin{array}{rcll} \dot{x} & = & Ax + B\mathrm{sat}(z+d_1) \ , & x(0)=0 \\ \dot{z} & = & -z + \epsilon(\exp(|x|+d_2)-1) \ , & z(0)=0 \end{array} \tag{66}$$

where $x \in \mathbb{R}^n$, $z \in \mathbb{R}$, the eigenvalues of A all have strictly negative real part, ϵ is a small parameter and $\mathrm{sat}(s) = \mathrm{sgn}(s)\min\{|s|,1\}$. This composite system is a well-defined interconnection of the subsystems

$$\Sigma_1 : \quad \begin{array}{rcll} \dot{x} & = & Ax + B\mathrm{sat}(u_1) \quad , & x(0)=0 \\ y_1 & = & |x| & \end{array} \tag{67}$$

and

$$\Sigma_2 : \quad \begin{array}{rcl} \dot{z} & = & -z + \epsilon\,(\exp(u_2) - 1) \quad , \qquad z(0) = 0 \\ y_2 & = & z \ . \end{array} \tag{68}$$

A gain function for the Σ_1 system is the product of the ∞-gain for the linear system

$$\begin{array}{rcl} \dot{x} & = & Ax + Bu \quad , \qquad x(0) = 0 \\ y & = & x \ , \end{array} \tag{69}$$

which we will call $\bar{\gamma}_1$, with the function sat(s), i.e. for the system Σ_1,

$$||y||_\infty \leq \bar{\gamma}_1 \text{sat}(||u_1||_\infty) \ . \tag{70}$$

For the system Σ_2, it is easy to see that

$$||z||_\infty \leq |\epsilon|\,(\exp(||u_2||_\infty) - 1) \ . \tag{71}$$

We must have that the distance between the curves $(s, \bar{\gamma}_1 \text{sat}(s))$ and $(|\epsilon|\,(\exp(r) - 1), r)$ grows without bound. Graphically, one can see that a necessary and sufficient condition for this to be the case is that

$$|\epsilon| < \frac{1}{\exp(\bar{\gamma}_1) - 1} \ . \tag{72}$$

3.4 General conic regions

There are many different ways to partition the ambient space in an attempt to establish the graph separation condition in (5). So far we have looked at only two very specific sufficient conditions, the small gain theorem and the passivity theorem. The general idea in these theorems is to constrain signals in the graph of Σ_1 to be inside some conic region, and signals in the inverse graph of Σ_2 to be outside of this conic region. Conic regions more general than those used for the small gain and passivity theorems can be generated by using operators on the input-output pairs of the feedback components.

Let $\mathbf{C}$ and $\mathbf{R}$ be operators on truncated ordered pairs in the ambient space and let γ be a gain function. We say that the graph of Σ_1 is **inside** CONE$(\mathbf{C}, \mathbf{R}, \gamma)$ if, for each $(u, y) =: z$ belonging to the graph of Σ_1,

$$||\mathbf{C}(z_\tau)|| \leq \gamma(||\mathbf{R}(z_\tau)||) \qquad \text{for all } \tau \ . \tag{73}$$

On the other hand, we say that the inverse graph of Σ_2 is **strictly outside** CONE$(\mathbf{C}, \mathbf{R}, \gamma)$ if there exists a proper separation function ρ such that, for each $(y, u) =: x$ belonging to the inverse graph of Σ_2,

$$||\mathbf{C}(x_\tau)|| \geq \gamma \circ (\text{Id} + \rho)(||\mathbf{R}(x_\tau)||) + \rho(||x_\tau||) \qquad \text{for all } \tau \ . \tag{74}$$

We will only consider the case where the maps $\mathbf{C}$ and $\mathbf{R}$ are **incrementally stable**, i.e. there exists a gain function $\bar{\gamma}$ such that, for each x_1 and x_2 in the ambient space and all τ,

$$\begin{aligned} \|\mathbf{C}(x_{1_\tau}) - \mathbf{C}(x_{2_\tau})\| &\leq \bar{\gamma}(\|x_{1_\tau} - x_{2_\tau}\|) \\ \|\mathbf{R}(x_{1_\tau}) - \mathbf{R}(x_{2_\tau})\| &\leq \bar{\gamma}(\|x_{1_\tau} - x_{2_\tau}\|) \ . \end{aligned} \tag{75}$$

In this case, the following result holds.

Nonlinear conic sector theorem: *If the graph of Σ_1 is inside* $\textsc{Cone}(\mathbf{C}, \mathbf{R}, \gamma)$ *and the inverse graph of Σ_2 is strictly outside* $\textsc{Cone}(\mathbf{C}, \mathbf{R}, \gamma)$ *then the well-defined interconnection is stable.*

In fact, when γ and ρ are linear functions, the well-defined interconnection is finite gain stable.

The small gain and passivity theorems we have discussed can be interpreted in the framework of the nonlinear conic sector theorem. For example, for the nonlinear small gain theorem, the operator $\mathbf{C}$ is a projection onto the second coordinate in the ambient space while $\mathbf{R}$ is a projection onto the first coordinate; γ is the gain function γ_1 and the small gain condition guarantees that the inverse graph of Σ_2 is strictly outside of the cone specified by this $\mathbf{C}$, $\mathbf{R}$ and γ.

In the remaining subsections we will discuss other useful choices for the operators $\mathbf{C}$ and $\mathbf{R}$.

3.4.1 The classical conic sector (circle) theorem

For linear SISO systems connected to memoryless nonlinearities, there is an additional classical result, known as the circle theorem, which follows from the nonlinear conic sector theorem using the 2-norm and taking

$$\begin{aligned} \mathbf{C}(u,y) &= y + cu \\ \mathbf{R}(u,y) &= ru \qquad r \geq 0 \\ \gamma(s) &= s \ . \end{aligned} \tag{76}$$

Indeed, suppose ϕ is a memoryless nonlinearity which satisfies

$$|\phi(u,t) + cu| \leq |ru| \qquad \text{for all } t, u \ . \tag{77}$$

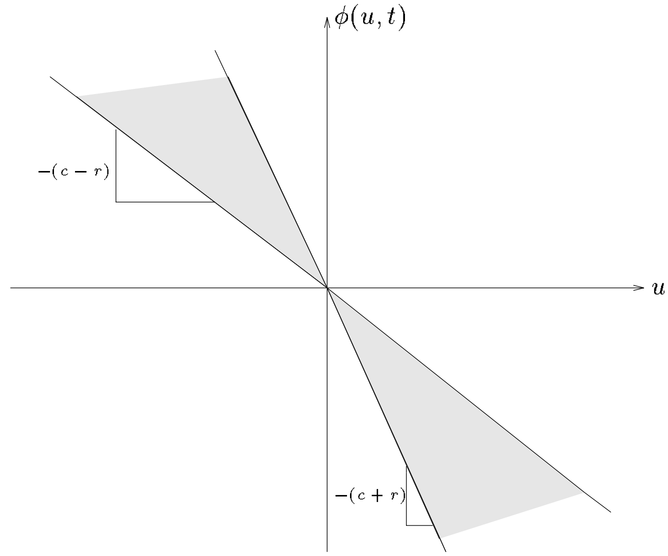

Figure 7: Instantaneous sector

Graphically, the constraint on ϕ is shown in figure 7. (In the case shown, $c > r > 0$.) We will use the notation $\textsc{Sector}[-(c+r), -(c-r)]$ for the memoryless nonlinearity. It is also clear that the graph of ϕ lies in the $\textsc{Cone}(\mathbf{C}, \mathbf{R}, \gamma)$ with $\mathbf{C}, \mathbf{R}, \gamma$ defined in (76). For a linear, time invariant, finite dimensional SISO system, whether its inverse graph is strictly outside of this cone can be determined by examining the Nyquist plot of its transfer function. The condition on the Nyquist plot is expressed relative to a disk $\mathcal{D}_{c,r}$ in the complex plane centered on the real axis passing through the points on the real axis with real part $-1/(c+r)$ and $-1/(c-r)$ as shown in figure 8.

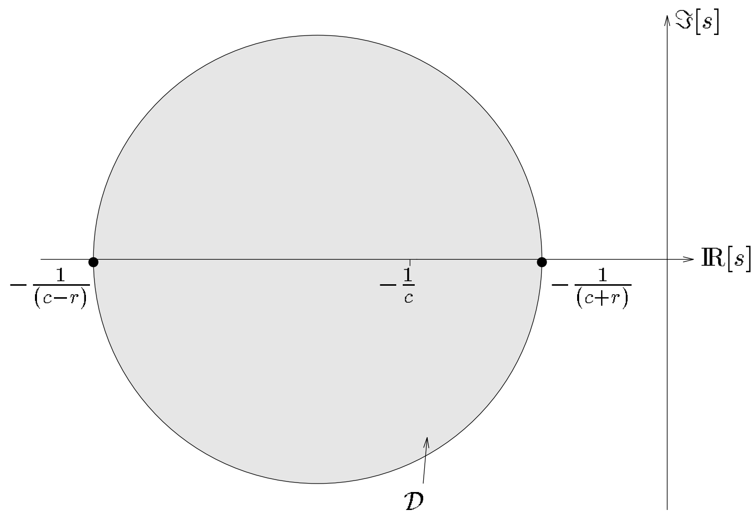

Figure 8: A disc in the complex plane

Circle theorem: *Let $r \geq 0$ and consider a well-defined interconnection of a memoryless nonlinearity belonging to* SECTOR$[-(c+r), -(c-r)]$ *with a SISO system having a real, rational transfer function $G(s)$. If:*

- *$r > c$, $G(s)$ is stable and the Nyquist plot of $G(s)$ lies in the interior of the disc $\mathcal{D}_{c,r}$, or*

- *$r = c$, $G(s)$ is stable and the Nyquist plot of $G(s)$ is bounded away and to the right of the vertical line passing through the real axis at the value $-1/(c+r)$, or*

- *$r < c$, the Nyquist plot of $G(s)$ (with Nyquist path indented into the right-half plane) is outside of and bounded away from the disc $\mathcal{D}_{c,r}$, and the number of times the plot encircles this disc in the counterclockwise direction is equal to the number of poles of $G(s)$ with strictly positive real part,*

then the interconnection is finite gain stable.

Case 1 is similar to the small gain theorem, while case 2 is similar to the passivity theorem. We will now explain case 3 in more detail. Let $n(s)$ and $d(s)$ represent, respectively, the numerator and denominator polynomials of $G(s)$. Since the point $(-1/c, 0)$ is inside the disc $\mathcal{D}_{c,r}$, it follows from the assumption of the theorem together with the well-known Nyquist stability condition that all of the roots of the polynomial $d(s) + cn(s)$ have negative real part. Then we can write $y = G(s)u = N(s)D(s)^{-1}u$ where

$$\begin{aligned} D(s) &:= \frac{d(s)}{d(s) + cn(s)} \\ N(s) &:= \frac{n(s)}{d(s) + cn(s)} \end{aligned} \tag{78}$$

and, by taking $z = D(s)^{-1}u$, we can describe all of the possible input-output pairs as

$$(u, y) = \begin{pmatrix} D(s)z, & N(s)z \end{pmatrix} . \tag{79}$$

Notice that $D(s) + cN(s) = 1$ so that

$$||u + cy||_2 = ||z||_2 . \tag{80}$$

To put a lower bound on this expression in terms of $||u||_2$ and $||y||_2$, for the purpose of showing that the graph is strictly outside of the cone defined in (76), we will need the 2-norm gains for systems modeled by the transfer functions $N(s)$ and $D(s)$. We will use the symbols γ_N and γ_D for these gains. The condition of the circle theorem guarantees that $\gamma_N < r^{-1}$. To see this, note that

$$N(s) = \frac{G(s)}{1 + cG(s)} \tag{81}$$

implying

$$\gamma_N := \sup_{\omega \in \mathbb{R}} \left| \frac{G(j\omega)}{1 + cG(j\omega)} \right| . \tag{82}$$

But

$$\begin{aligned}|1+c\,G(j\omega)|^2-r^2|G(j\omega)|^2 &= (c\,\mathrm{Re}\,\{G(j\omega)\}+1)^2+c^2\,\mathrm{Im}^2\,\{G(j\omega)\}\\ &\quad -r^2\,\mathrm{Re}^2\,\{G(j\omega)\}-r^2\,\mathrm{Im}^2\,\{G(j\omega)\}\\ &= (c^2-r^2)\left(\mathrm{Re}\,\{G(j\omega)\}+\frac{c}{c^2-r^2}\right)^2\\ &\quad +(c^2-r^2)\,\mathrm{Im}^2\,\{G(j\omega)\}-\frac{r^2}{c^2-r^2}\;.\end{aligned} \tag{83}$$

Setting the latter expression to zero defines the boundary of the disc $\mathcal{D}_{c,r}$. Since outside of this disc the expression is positive, it follows that $\gamma_N < r^{-1}$.

Returning to the calculation initiated in (80), note that $\gamma_N < r^{-1}$ implies that there exists a strictly positive real number ϵ such that

$$(1-\epsilon\gamma_D)\gamma_N^{-1} \geq r+2\epsilon\;. \tag{84}$$

So,

$$\begin{aligned}||u+cy||_2 &= ||z||_2=(1-\epsilon\gamma_D)||z||_2+\epsilon\gamma_D||z||_2\\ &\geq (1-\epsilon\gamma_D)\gamma_N^{-1}||y||_2+\epsilon||u||_2\\ &\geq (r+\epsilon)||y||_2+\epsilon(||u||_2+||y||_2)\;.\end{aligned} \tag{85}$$

We conclude that the inverse graph of the linear system is strictly outside of the $\textsc{Cone}(\mathbf{C},\mathbf{R},\gamma)$ as defined in (76).

Note, incidentally, that $N(s)$ is the closed loop transfer function from d_1 to y_1 for the special case where the memoryless nonlinearity satisfies $\phi(u)=-cu$. This suggests another way of determining stability: first make a preliminary loop transformation with the feedback $-cu$, changing the original linear system into the system with transfer function $N(s)$ and changing the nonlinearity into a new nonlinearity $\tilde{\phi}$ satisfying $|\tilde{\phi}(u,t)| \leq r|u|$, and then apply the classical small gain theorem to the resulting feedback system.

Example 3.5 Let

$$G(s)=\frac{175}{(s-1)(s+4)^2}\;. \tag{86}$$

The Nyquist plot of $G(s)$ is shown in figure 9. Since G(s) has one pole with positive real part only the third condition of the circle theorem can apply. A disc centered at -8.1

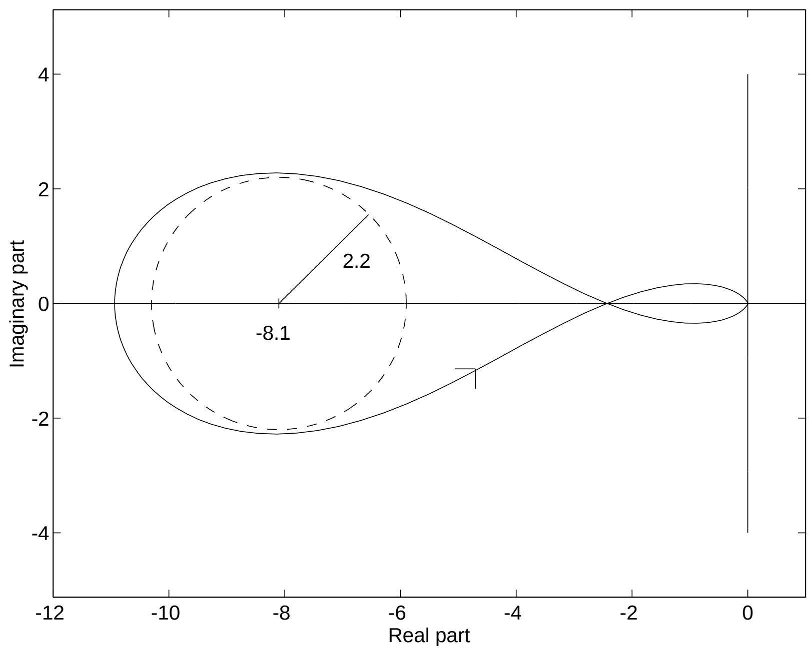

Figure 9: The Nyquist plot for $G(s)$ in example 3.5.

on the real axis and with radius 2.2 can be placed inside the left loop of the Nyquist plot. Such a disc corresponds to the values $c = 0.293$ and $r = 0.079$. Since the Nyquist plot encircles this disc once in the counterclockwise direction, it follows that the standard feedback connection with the feedback components $G(s)$ and a memoryless nonlinearity constrained to the SECTOR[-0.372,-0.214] is stable using the 2-norm.

3.4.2 Coprime fractions

Typical input-output stability results based on stable coprime fractions are corollaries of the conic sector theorem. For example, suppose both Σ_1 and Σ_2 are modeled by transfer functions $G_1(s)$ and $G_2(s)$. Moreover, assume there exist stable (in any p-norm) transfer functions N_1, D_1, $\tilde{N}_1, \tilde{D}_1$, N_2 and

D_2 such that D_1, D_2 and $\tilde{D}_1$ are invertible and

$$\begin{aligned} G_1 &= N_1 D_1^{-1} = \tilde{D}_1^{-1}\tilde{N}_1 \\ G_2 &= N_2 D_2^{-1} \\ \mathrm{Id} &= \tilde{D}_1 D_2 - \tilde{N}_1 N_2 \ . \end{aligned} \tag{87}$$

Let $\mathbf{C}(u,y) = \tilde{D}_1(s)y - \tilde{N}_1(s)u$, which is incrementally stable in any p-norm, let $\mathbf{R}(u,y) = 0$ and let $\gamma \equiv 0$. Then, it turns out that the graph of Σ_1 is inside and the inverse graph of Σ_2 is strictly outside $\textsc{Cone}(\mathbf{C}, \mathbf{R}, \gamma)$ and thus the feedback loop is finite gain stable in any p-norm.

To verify these claims about the properties of the graphs, first recognize that the graph of Σ_i can be represented as

$$\mathcal{G}_{\Sigma_i} = \left(\ D_i(s)z \ , \quad N_i(s)z \ \right) \tag{88}$$

where z represents any reasonable signal. Then, for signals in the graph of Σ_1,

$$\mathbf{C}\left(D_1(s)z_\tau, N_1(s)z_\tau\right) = \tilde{D}_1(s)N_1(s)z_\tau - \tilde{N}_1(s)D_1(s)z_\tau \equiv 0 \ . \tag{89}$$

Conversely, for signals in the inverse graph of Σ_2,

$$\begin{aligned} \left\| \mathbf{C}\left(\ N_2(s)z_\tau \ , \ D_2(s)z_\tau \ \right) \right\| &= \left\| \tilde{D}_1(s)D_2(s)z_\tau - \tilde{N}_1(s)N_2(s)z_\tau \right\| \\ &= \|z_\tau\| \\ &\geq \epsilon \left\| \left(\ N_2(s)z_\tau \ , \ D_2(s)z_\tau \ \right) \right\| \end{aligned} \tag{90}$$

for some strictly positive real number ϵ. The last inequality follows from the fact that D_2 and N_2 are finite gain stable.

Example 3.6 (This example is drawn from the work in Potvin, M-J., J. Jeswiet, and J.-C. Piedboeuf, *Trans. of NAMRI/SME.* vol. XXII, pp. 373-377, 1994.) Let Σ_1 represent the fractional Voigt-Kelvin model for the relation between stress and strain in structures displaying plasticity. For suitable values of Young's modulus, damping magnitude, and order of derivative for the strain, the transfer function of Σ_1 is given by

$$g_1(s) = \frac{1}{1+\sqrt{s}}.$$

Integral feedback control, $g_2(s) = -\frac{1}{s}$, may be used for purposes of asymptotic tracking. Here we can take

$$\begin{array}{ll} N_1(s) = \dfrac{1}{s+1} & D_1(s) = \dfrac{1+\sqrt{s}}{s+1} \\ N_2(s) = -\dfrac{s+1}{1+s(1+\sqrt{s})} & D_2(s) = \dfrac{s(s+1)}{1+s(1+\sqrt{s})} \end{array} \tag{91}$$

It can be shown that these fractions are stable linear operators, and thereby incrementally stable, in the 2-norm. (This fact, in essence, is equivalent to proving nominal stability and can be shown using Nyquist theory.) Moreover, it is easy to see that $D_1D_2 - N_1N_2 = 1$ so that the feedback loop is stable, in fact finite gain stable.

3.4.3 Robustness of stability and the gap metric

It is clear from the original graph separation theorem that if a well-defined interconnection is stable, i.e. the appropriate graphs are separated in distance, then modifications of the feedback components will not destroy stability if the modified graphs are close to the original graphs.

Given two systems Σ_1 and Σ define $\vec{\delta}(\Sigma_1, \Sigma) = \alpha$ if α is the smallest number for which

$$x \in \mathcal{G}_\Sigma, \quad \Longrightarrow \quad d_\tau(x, \mathcal{G}_{\Sigma_1}) \leq \alpha \|x\|_\tau \quad \text{for all } \tau \ .$$

The quantity $\vec{\delta}(\cdot,\cdot)$ is called the "directed gap" between the two systems and characterizes basic neighborhoods where stability as well as closed loop properties are preserved under small perturbations from the nominal system Σ_1 to a nearby system Σ.

More specifically, if the interconnection of (Σ_1, Σ_2) is finite gain stable, we define the gain $\beta_{\Sigma_1,\Sigma_2}$ to be the smallest real number such that

$$\|\begin{pmatrix} d_1 + y_2 \\ y_1 \end{pmatrix}\|_\tau \leq \beta_{\Sigma_1,\Sigma_2} \|\begin{pmatrix} d_1 \\ d_2 \end{pmatrix}\|_\tau, \quad \text{for all } \tau \ .$$

If Σ is such that

$$\vec{\delta}(\Sigma_1, \Sigma)\beta_{\Sigma_1,\Sigma_2} < 1 \ ,$$

then the interconnection of (Σ, Σ_2) is also finite gain stable.

As a special case let $\Sigma, \Sigma_1, \Sigma_2$ represent linear systems acting on finite energy signals. Further, assume that there exist stable transfer functions N, D where D is invertible, $G_1 = ND^{-1}$ and normalized so that

$D^T(-s)D(s) + N^T(-s)N(s) = \mathrm{Id}$. Then, the class of systems in a ball with radius $\gamma \geq 0$, measured in the directed gap, is given by $\text{CONE}(\mathbf{C}, \mathbf{R}, \gamma)$, where $\mathbf{R} = \mathrm{Id}$ and

$$\mathbf{C} = \mathrm{Id} - \begin{pmatrix} D(s) \\ N(s) \end{pmatrix} \mathbf{\Pi}_+(D^T(-s), N^T(-s))$$

where $\mathbf{\Pi}_+$ designates truncation part of the Laplace transform of finite energy signals to the part with poles in the left half plane. At the same time, if $\beta_{\Sigma_1,\Sigma_2} < 1/\gamma$, then it can be shown that Σ_2 is strictly outside the cone $\text{CONE}(\mathbf{C}, \mathbf{R}, \gamma)$ and, therefore, stability of the interconnection of Σ with Σ_2 is guaranteed for any Σ inside $\text{CONE}(\mathbf{C}, \mathbf{R}, \gamma)$.

Given Σ and Σ_1, the computation of the gap reduces to a standard $\mathcal{H}_\infty$-optimization problem (see [Georgiou and Smith, 1990]). Also, given Σ_1, the computation of a controller Σ_2 which stabilizes a maximal cone around Σ_1 reduces to a standard $\mathcal{H}_\infty$-optimization problem ([Georgiou and Smith, 1990]) and forms the basis of the $\mathcal{H}_\infty$-loop shaping procedure for linear systems introduced in [McFarlane and Glover, 1989].

A second key result which motivated introducing the gap metric is the claim that the behavior of the feedback interconnection of Σ and Σ_2 is "similar" to that of the interconnection of Σ_1 and Σ_2 if and only if the distance between Σ and Σ_1, measured using the gap metric, is small (i.e., Σ lies within a "small aperture" cone around Σ_1). The "gap" function is defined as

$$\delta(\Sigma_1, \Sigma) = \max\{\vec{\delta}(\Sigma_1, \Sigma), \vec{\delta}(\Sigma, \Sigma_1)\}$$

to "symmetrize" the distance function $\vec{\delta}(\cdot, \cdot)$ with respect to the order of the arguments. Then, the above claim can be stated more precisely as follows: for each $\epsilon > 0$ there exists a $\zeta(\epsilon) > 0$ such that

$$\delta(\Sigma_1, \Sigma) < \zeta(\epsilon) \Longrightarrow \|x - x_1\|_\tau < \epsilon \|d\|_\tau$$

where $d = \binom{d_1}{d_2}$ is an arbitrary signal in the ambient space and x (resp. x_1) represents the response $\binom{d_1+y_2}{y_1}$ of the feedback interconnection of (Σ, Σ_2) (resp. (Σ_1, Σ_2)). Conversely, if $\|x - x_1\|_\tau < \epsilon \|d\|_\tau$ for all d and τ, then $\delta(\Sigma_1, \Sigma) \leq \epsilon$.

4 Defining terms

Ambient space: The Cartesian product space containing the inverse graph of Σ_2 and the graph of Σ_1.

Distance (from a signal to a set): Measured using a norm function. The infimum, over all signals in the set, of the norm of the difference between the signal and a signal in the set. See equation (4). Used to characterize necessary and sufficient conditions for input-output stability. See section 2.

Dynamical system: An object which produces an output signal for each input signal.

Feedback components: The dynamical systems which make up a well-defined interconnection.

Finite gain stable system: A dynamical system is finite gain stable if there exists a nonnegative constant such that, for each input-output pair, the norm of the output is bounded by the norm of the input times the constant.

Gain function: A function from the nonnegative real numbers to the nonnegative real numbers which is continuous, nondecreasing and zero when its argument is zero. Used to characterize stability. See section 2. (Some form of) the symbol γ is usually used.

Graph (of a dynamical system): The set of ordered input-output pairs (u, y).

Inner product: Defined for signals of the same dimension defined on the semi-infinite interval. The integral from zero to infinity of the component-wise product of the two signals.

Inside (or strictly outside) $\textsc{Cone}(\mathbf{C}, \mathbf{R}, \gamma)$: Used to characterize the graph or inverse graph of a system. Determined by whether or not signals in the graph or inverse graph satisfy certain inequalities involving the operators $\mathbf{C}$ and $\mathbf{R}$ and the gain function γ. See equations (73) and (74). Used in the conic sector theorem. See section 3.4.

Inverse graph (of a dynamical system): The set of ordered output-input pairs (y, u).

Norm function ($\|\cdot\|$): Used to measure the size of signals defined on the semi-infinite interval. Examples are the p-norms $p \in [1, \infty]$ (see equations (1) and (2)).

Parseval's theorem: Used to make connections between properties of graphs for SISO systems modeled with real, rational transfer functions and frequency domain characteristics of their transfer functions. Parseval's theorem relates the inner product of signals to their Fourier transforms if they exist. For example, it states that, if two scalar signals u and y, assumed to be zero for negative values of time, have Fourier transforms $\hat{u}(j\omega)$ and $\hat{y}(j\omega)$ then

$$\langle u, y \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{y}^*(j\omega)\hat{u}(j\omega) d\omega \ .$$

Passive: Terminology motivated from electrical network theory. A dynamical system is passive if the inner product of each input-output pair is nonnegative.

Proper separation function: A function from the nonnegative real numbers to the nonnegative real numbers which is continuous, zero at zero, strictly increasing and unbounded. Such functions are invertible on the nonnegative real numbers. Used to characterize nonlinear separation theorems. See section 3.3. (Some form of) the symbol ρ is usually used.

Semi-infinite interval: The time interval $[0, \infty)$.

Signal: A "reasonable" vector-valued function defined on a finite or semi-infinite time interval. By "reasonable" we mean piecewise continuous or measurable.

SISO systems: An abbreviation for single input, single output systems.

Stable system: A dynamical system is stable if there exists a gain function such that, for each input-output pair, the norm of the output is bounded by the gain function evaluated at the norm of the input.

Strict contraction: The composition of two gain functions γ_1 and γ_2 is a strict contraction if there exists a proper separation function ρ such that $(\gamma_1 + \rho) \circ (\gamma_2 + \rho) \leq \mathrm{Id}$. Recall that $\mathrm{Id}(s) = s$ and $\tilde{\gamma}_1 \circ \tilde{\gamma}_2(s) = \tilde{\gamma}_1(\tilde{\gamma}_2(s))$. Graphically, this is the equivalent to the curve $(s, \gamma_1(s) + \rho(s))$ never being above the curve $(\gamma_2(r) + \rho(r), r)$. This concept is used to state the nonlinear small gain theorem. See section 3.3.2.

Strictly passive: We have used various notions of strictly passive including input-, output-, input and output-, and nonlinear input and output- strictly passive. All notions strengthen the requirement that the inner product of the input-output pairs be positive by requiring a positive lower bound that depends on the 2-norm of the input and/or output. See section 3.2.

Truncated signal: A signal defined on the semi-infinite interval which is derived from another signal (not necessarily defined on the semi-infinite interval) by first appending zeros to extend the signal onto the semi-infinite interval and then keeping the first part of the signal and setting the rest of the signal to zero. Used to measure the size of finite portions of signals.

Well-defined interconnection: An interconnection of two dynamical systems in the configuration of figure 1 which results in another dynamical system, i.e. one in which an output signal is produced for each input signal.

5 References

Georgiou, T.T. and M.C. Smith, 1990. Optimal robustness in the gap metric. *IEEE Trans. on Automatic Control*, 35:673–686.

Jiang, Z.P., A.R. Teel and L. Praly, 1995. Small-gain theorem for ISS systems and applications. *Mathematics of Control, Signals, and Systems*, 7(2):95–120.

Mareels, I.M.Y. and D.J. Hill, 1992. Monotone stability of nonlinear feedback systems. *Journal of Mathematical Systems, Estimation and Control*, 2(3):275–291.

McFarlane, D.C. and K. Glover, 1989. *Robust controller design using normalized coprime factor plant descriptions.* Lecture Notes in Control and Information Sciences, Springer-Verlag, vol. 138.

Safonov, M., 1980. *Stability and robustness of multivariable feedback systems.* The MIT Press, Cambridge, MA.

Sandberg, I.W., 1964. On the $\mathcal{L}_2$-boundedness of solutions of nonlinear functional equations. *Bell Sys. Tech. J.*, 43:1581–1599.

Sontag, E., 1989. Smooth stabilization implies coprime factorization. *IEEE Transactions on Automatic Control*, 34:435–443.

Zames, G., 1966a. On the input-output stability of time-varying nonlinear feedback systems. Part I: Conditions using concepts of loop gain, conicity, and positivity. *IEEE Transactions on Automatic Control*, 11:228–238.

G. Zames, 1966b. On the input-output stability of time-varying nonlinear feedback systems. Part II: Conditions involving circles in the frequency plane and sector nonlinearities. *IEEE Transactions on Automatic Control*, 11:465–476.

6 For further information

As mentioned at the outset, the material presented in this chapter is based on the results in [Sandberg, 1964], [Zames, 1966(a,b)] and [Safonov, 1980]. In the latter, a more general feedback interconnection structure is considered where nonzero initial conditions can also be consider as inputs.

Other excellent references on input-output stability include *The Analysis of Feedback Systems*, 1971, by J.C. Willems and *Feedback systems: input-output properties*, 1975, by C. Desoer and M. Vidyasagar. A nice text addressing the factorization method in linear systems control design is *Control Systems Synthesis: a Factorization Approach*, 1985, by M. Vidyasagar. A treatment of input-output stability for linear, infinite dimensional systems can be found in chapter 6 of *Nonlinear Systems Analysis*, 1993, by M. Vidyasagar. That chapter also discusses many of the connections between input-output stability and state-space (Lyapunov) stability. Another excellent reference is *Nonlinear Systems*, 1992, by H. Khalil.

There are results similar to the circle theorem that we have not discussed. These results go under the heading of "multiplier" results and apply to feedback loops with a linear element and a memoryless, nonlinear element with extra restrictions such as time invariance and constrained slope. Special cases are the well-known Popov and off-axis circle criterion. Some of these results can be recovered using the general conic sector theorem although we have not taken the time to do this. Other results, like the Popov criterion, impose extra smoothness conditions on the external inputs which are not found in the standard problem. References for these problems are *Hyperstability of Control Systems*, 1973, V.M. Popov and *Frequency domain criteria for absolute stability*, 1973, by K.S. Narendra and J.H. Taylor.

Another topic closely related to these multiplier results is the structured small gain theorem for linear systems which motivates much of the μ-synthesis control design methodology. This is described, for example, in *μ-Analysis and synthesis toolbox*, 1991, by G. Balas, J. Doyle, K. Glover, A. Packard and R. Smith.

There are many advanced topics concerning input-output stability that we have not addressed. These include the study of small-signal stability, well-posedness of feedback loops, and control design based on input-output stability principles. Many articles on these topics frequently appear in control and systems theory journals such as *IEEE Transactions on Automatic Control*, *Automatica*, *International Journal of Control*, *Systems and Control Letters*, *Mathematics of Control, Signals, and Systems*, to name a few.