Control of Nonlinear Systems.

S. T. Glad
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden

October 27, 2009

.

Contents

Chapter 1

Introduction.

1.1 Some examples of nonlinear systems

Our subject is the control of nonlinear systems. To get a feeling for it, let us consider some examples.

Example 1.1 Consider a very simplified model for velocity control of an aircraft. If the velocity is x_1 and the mass normalized to 1, then

$$\dot{x}_1 = x_2 - f(x_1) \tag{1.1}$$

where x_2 is the engine thrust and $f(x_1)$ is the aerodynamic drag. A simplified engine model is just a time constant from pilot command u to engine thrust:

$$\dot{x}_2 = -x_2 + u \tag{1.2}$$

Together (1.1) and (1.2) form a model of the aircraft velocity control. ■

Example 1.2 Consider the heat exchanger described in figure 1.1. A fluid which initially has the temperature T_0 flows with the flow rate q through the heat exchanger, which is surrounded by a medium with temperature T_h. It is assumed that very good mixing takes place so that one can assume the same temperature T at every point in the heat exchanger. If the heat capacity of the fluid is c per unit volume and C for the whole heat exchanger, and if the heat transfer coefficient of the walls is κ, then a heat balance gives

$$\frac{d}{dt}(CT) = qcT_0 - qcT + \kappa(T_h - T)$$

Assume that the flow rate is controlled around a nominal flow q_0 so that

$$q = q_0 - u$$

Then, using the numerical values

$$c/C = 1, \;\; \kappa/C = 1, \;\; T_h = q_0 = -T_0 = 1$$

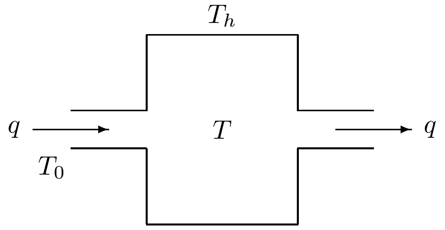


Figure 1.1: A simple heat exchanger model

gives the model

$$\dot{T} = -2T + uT + u \tag{1.3}$$

where the temperature T is a state variable and the flow change u is the input. (Note that a positive u means a decrease in flow.) ■

Example 1.3 A pendulum where the length d is varied, is described by

$$\ddot{\theta} + 2\dot{d}\dot{\theta}/d + (g/d)\sin\theta = 0$$

Defining $x_1 = \theta$ and $x_2 = \dot{\theta}$ we get the equations

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= (-2x_2\dot{d} - g\sin x_1)/d \end{aligned} \tag{1.4}$$

■

Example 1.4 Consider a rigid body rotating freely in space. Let x_i be the angular velocity along the i:th principal axis. Let the external torque be a vector whose components along the principal axes are u_1, u_2 and u_3. The equations are then

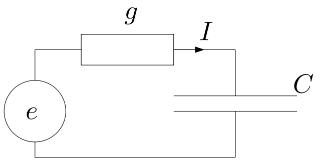
$$\begin{aligned} \dot{x}_1 &= a_1x_2x_3 + u_1 \\ \dot{x}_2 &= a_2x_1x_3 + u_2 \\ \dot{x}_3 &= a_3x_1x_2 + u_3 \end{aligned} \tag{1.5}$$

where

$$a_1 = \frac{I_2 - I_3}{I_1}, \quad a_2 = \frac{I_3 - I_1}{I_2}, \quad a_3 = \frac{I_1 - I_2}{I_3}$$

with the I_i being the moments of inertia. Here there are three state variables (x_1, x_2 and x_3) and three input variables (u_1, u_2 and u_3). ■

Example 1.5 Consider the following electrical circuit



The resistive element is assumed to have a voltage drop $g(I)$ which is a possibly nonlinear function of the current. If the voltage across the capacitor is v, the system is described by the equations

$$\begin{aligned} C\dot{v} &= I \\ e &= v + g(I) \end{aligned} \tag{1.6}$$

If the function g has an inverse g^{-1} we get the following description

$$\dot{v} = g^{-1}(e - v)/C \tag{1.7}$$

■

1.2 Discussion of the examples

Is there a common description for all the examples? We see that (1.1,1.2), (1.3), (1.4), (1.5) and (1.7) are all special cases of the description

$$\dot{x} = f(x, u, \dot{u}, \dots, u^{(\alpha)}) \tag{1.8}$$

where x is a vector of *state* variables and u is a vector values external signal, often called *input* signal or *control* signal. If we exclude (1.4) we can write

$$\dot{x} = f(x, u) \tag{1.9}$$

This is often considered to be the standard description of a nonlinear system. If we also exclude (1.7) then all the examples can be written in the form

$$\dot{x} = f(x) + g(x) \cdot u \tag{1.10}$$

where the right hand side is an affine function of u.

It should be noted that many control problems, which are not of the form (1.10), can be put into that form by the introduction of extra variables and a redefinition of the input. In (1.4) for instance we could regard d as a state and $\dot{d}$ as an input: $x_3 = d$, $u = \dot{d}$ giving the description

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= (-2x_2u - g\sin x_1)/x_3 \\ \dot{x}_3 &= u \end{aligned} \tag{1.11}$$

which is of the form (1.10). Similarly in (1.7) we could regard the derivative of the voltage e as the input giving

$$\begin{aligned} \dot{v} &= g^{-1}(e - v)/C \\ \dot{e} &= u \end{aligned} \tag{1.12}$$

which is also of the form (1.10).

Note that the system (1.3) is a very special case of the system description (1.10), since f is linear and g is affine. Such a system is called *bilinear*. The general form is

$$\dot{x} = Ax + \sum_{j=1}^{m} u_j D_j x + Bu \tag{1.13}$$

where A and D_j are $n \times n$ matrices and B is an $n \times m$ matrix (where n is the dimension of x and m is the dimension of u).

Finally we should note that the most general description we have discussed so far, (1.8) is not general enough to cover all physical descriptions that might be of interest. The description (1.7), for instance, was arrived at using the assumption that the function g is invertible. If that is not the case, then we are stuck with the description (1.6) which is a combination of algebraic and differential equations. Therefore it would be more natural to regard descriptions of the form

$$g_i(x, \dot{x}, \dots, x^{(r)}, u, \dot{u}, \dots, u^{(\alpha)}) = 0, \quad i = 1, \dots, N \tag{1.14}$$

as standard descriptions. The theory for such systems is however in the early stages of development. Therefore we will concentrate on system descriptions of the form (1.8), (1.9) and (1.10).

1.3 Exercises.

1.1 Show that the description (1.8) is no more general than (1.9) for linear systems, i. e. show that there exists a transformation

$$z = Tx + \sum S_i u^{(i)}$$

with nonsingular T, which transforms

$$\dot{x} = Ax + B_0 u + B_1 \dot{u} + \cdots + B_r u^{(r)}$$

into the standard description

$$\dot{z} = Fz + Gu$$

Chapter 2

Nonlinear differential equations.

In the previous chapter we discussed the desription of nonlinear systems and arrived at the equation

$$\dot{x} = f(x, u, \dot{u}, \ldots, u^{(\alpha)}) \tag{2.1}$$

and its sligthly more specialized form

$$\dot{x} = f(x, u) \tag{2.2}$$

In this chapter some basic properties of differential equations are discussed. The first and obvious question is: Suppose we put a continuous input signal $u = u(t)$ into the right hand side. Can we then be sure that there exists a solution of the differential equation? If we fix the time function $u(t)$, then both (2.1) and (2.2) can be seen as special cases of a differential equation with time varying right side:

$$\dot{x} = f(t, x)$$

2.1 Existence and uniqueness of solutions.

Consider a nonlinear differential equation

$$\begin{aligned} \dot{x} &= f(t, x) \\ x(t_0) &= x_0 \end{aligned} \tag{2.3}$$

where x is an n-vector, t is a real number (the time in almost all control theory applications) and where f is a continuous function. It turns out that it is more convenient to analyze the equivalent integral equation

$$x(t) = x_0 + \int_{t_0}^{t} f\big(\tau, x(\tau)\big) d\tau \tag{2.4}$$

Initially only the local properties close to the starting point of (2.3) are investigated. Consider therefore t- and x-values satisfying

$$t_0 \le t \le t_0 + a, \quad |x - x_0| \le b \tag{2.5}$$

Assume that for all t and x satisfying (2.5) it is true that

$$|f(t,x)| \le ba^{-1} \tag{2.6}$$

$$|f(t,x_1) - f(t,x_2)| \le \Lambda |x_1 - x_2| \tag{2.7}$$

and that

$$\theta = a\Lambda < 1 \tag{2.8}$$

In (2.6) and (2.7) the vertical bars denote the Euclidian vector norm. The inequality (2.7) is usually called a *Lipschitz condition* on f.

Remark 2.1 Obviously (2.6) and (2.8) can always be satisfied if a is chosen small enough, i.e. if the time interval is small enough.

For a continuous function v we define

$$||v|| = \max_{t_0 \le t \le t_0 + a} |v(t)| \tag{2.9}$$

We can now state the following existence and uniqueness result.

Theorem 2.1 A differential equation (2.3) which satisfies the conditions (2.5) - (2.8) has a unique solution on the interval $t_0 \le t \le t_0 + a$.

Proof. We will construct the solution, using the iteration

$$x_{n+1} = x_0 + \int_{t_0}^{t} f\big(\tau, x_n(\tau)\big) d\tau$$

and starting with the constant function

$$x_0(t) = x_0$$

a) We show that

$$|x_n(t) - x_0| \le b; \quad t_0 \le t \le t_0 + a$$

Obviously this is true for $n = 0$. Suppose it is known for all integers up to n. Then

$$|x_{n+1}(t) - x_0| \le \int_{t_0}^{t} |f(\tau, x_n(\tau))| d\tau \le \frac{b}{a} \int_{t_0}^{t} d\tau \le b$$

The result follows from induction.

b) We show that x_n converges to a limit. Having shown a) we know that we can apply (2.6) - (2.8) to all the x_n. Now consider the difference between two iterates

$$|x_{n+1}(t) - x_n(t)| \le \int_{t_0}^{t} |f(\tau, x_n(\tau)) - f(\tau, x_{n-1}(\tau))| \quad d\tau \le$$

$$\le \Lambda \int_{t_0}^{t} |x_n(\tau) - x_{n-1}(\tau)| \quad d\tau \le a\Lambda ||x_n - x_{n-1}|| = \theta ||x_n - x_{n-1}|| \tag{2.10}$$

Using this estimate repeatedly we get

$$||x_{n+1} - x_n|| \leq \theta ||x_n - x_{n-1}|| \leq \ldots \leq \theta^n ||x_1 - x_0||$$

If $m > n$ then

$$||x_m - x_n|| \leq ||x_m - x_{m-1}|| + \cdots + ||x_{n+1} - x_n|| \leq (\theta^{m-1} + \ldots + \theta^n)\, ||x_1 - x_0|| \leq$$

$$\leq \frac{\theta^n}{1-\theta} ||x_1 - x_0||$$

This expression converges to zero as n goes to infinity and $\{x_n\}$ is thus a Cauchy sequence. In particular, $x_n(t)$, for fixed t, is a Cauchy sequence of real numbers. It then has to converge to some value $x(t)$. Since this holds for all t in the chosen interval, we have shown that

$$x_n(t) \to x(t), \quad t_0 \leq t \leq t_0 + a,$$

for some function $x(t)$.

c) Show that x is continuous and satisfies (2.4). Since

$$|x(t+h) - x(t)| \leq |x(t+h) - x_n(t+h)| + |x_n(t+h) - x_n(t)| + |x_n(t) - x(t)| \leq$$

$$\leq 2||x - x_n|| + |x_n(t+h) - x_n(t)|$$

and each x_n is continuous, it follows that x is a continuous function.

Consider

$$|x_n(t) - x_0 - \int_{t_0}^{t} f\big(\tau, x(\tau)\big) d\tau| \leq \int_{t_0}^{t} |f(\tau, x_{n-1}(\tau)) - f(\tau, x(\tau))| \quad d\tau \leq \theta ||x_{n-1} - x||$$

It follows that

$$x_n(t) \to x_0 + \int_{t_0}^{t} f\big(\tau, x(\tau)\big) d\tau$$

as $n \to \infty$. As $x_n \to x$ it follows that x satisfies (2.4).

d) Show that x is a unique solution. Suppose there are two solutions x and z. Then using the same reasoning as in (2.10),

$$||x - z|| \leq \theta ||x - z||$$

Since $\theta < 1$, this implies that $||x - z|| = 0$ and consequently that $x = z$. ■

Remark 2.2 If f is continuous but does not satisfy the Lipschitz condition (2.7), then one can still prove existence but the solution is not necessarily unique, as shown by the differential equation

$$\frac{d}{dt} x = \sqrt{x}, \quad x(0) = 0$$

which has the solutions

$$x = 0, \quad x = \frac{t^2}{4}$$

Remark 2.3 Theorem 2.1 guarantees only local existence, since the time interval might have to be chosen small as explained in Remark 2.1. If a is too large there might not exist a solution over the whole time interval as shown by the differential equation.

$$\frac{d}{dt}x = x^2, \quad x(0) = 1$$

The solution is

$$x = \frac{1}{1-t}$$

which only exists for $t < 1$.

The phenomenon of Remark 2.3 is called "explosion" or "finite escape time". An interesting fact is that this is the only way in which global existence can be lost. This is formalized by the following theorem.

Theorem 2.2 Let $f(t,x)$ be continuous and satisfy (2.7) in the set

$$M = \{(t,x) : t_0 \leq t \leq t_1, \quad |x| \leq A\}$$

Let the starting point satisfy $|x_0| \leq A$. Then either there is a solution defined on the whole time interval $t_0 \leq t \leq t_1$ or else there is a solution on $t_0 \leq t \leq t_e, \quad (t_e < t_2)$, with $|x(t_e)| = A$. In other words the solution leaves the set M at $t = t_e$.

Remark 2.4 Often it is possible to give an upper bound on the solution x , showing that it can not leave the set M of Theorem 2.2. It then follows that there exists a solution on the whole time interval.

2.2 Continuity and differentiability with respect to initial values and parameters.

We will now consider a whole family of solutions with different initial conditions. We write

$$\begin{array}{rcl} \dot{x} & = & f(x) \\ x(0) & = & y \end{array} \tag{2.11}$$

To simplify the notation, we have assumed that f does not depend explicitly on t. Then there is no loss in generality in assuming that the initial time is $t = 0$. The results of this section are easily extended to the case where f is time dependent however.

Our task will be to find out how the solutions vary when the initial condition y is varied. To do that, the following lemma is needed.

Lemma 2.1 (*Gronwall's lemma*) If the continuous and nonnegative functions m and g satisfy

$$m(t) \leq C + \int_0^t g(\tau)m(\tau)\, d\tau, \quad 0 \leq t \leq T \tag{2.12}$$

for a positive constant C, then

$$m(t) \leq C \exp\left(\int_0^t g(\tau) d\tau\right), \quad 0 \leq t \leq T \tag{2.13}$$

Proof. Define

$$h(t) = C + \int_0^t g(\tau)m(\tau)\, d\tau$$

Differentiating gives

$$\dot{h}(t) = m(t)g(t) \leq h(t)g(t)$$

showing that

$$\frac{\dot{h}(t)}{h(t)} \leq g(t)$$

Integrating both sides then gives the desired result. ■

We now define the function

$$F(t, y)$$

as the solution of (2.11) at time t. The solution is thus regarded as a function of two variables: the time t and the initial state y. First it will be shown that the solution depends continuously (in fact Lipschitz continuously) on the initial condition.

Theorem 2.3 Let the differential equation (2.11) and the point x_0 be given. Asume that the conditions (2.6)-(2.8) are satisfied. Then there is a neighborhood V of x_0 and an $\epsilon > 0$ such that for every $y \in V$ there is a unique solution of (2.11) on $[0, \epsilon]$. Furthermore

$$|F(t, z) - F(t, y)| \leq e^{\Lambda t}|z - y| \tag{2.14}$$

Proof. The first part of the lemma follows directly from the proof of the existence theorem. To show (2.14) define

$$\phi(t) = |F(t, z) - F(t, y)|$$

Then

$$\phi(t) = |z - y + \int_0^t (f(F(s, z)) - f(F(s, y)))ds| \leq |z - y| + \Lambda \int_0^t \phi(s)ds$$

Using Gronwall's lemma immediately gives

$$\phi(t) \leq e^{\Lambda t}|z - y|$$

■

We can now go one step further and ask ourselves if the solution can be differentiated with respect to the initial condition. The following notation will be used. For a vector valued function $f(x)$, $f_x(x)$ will denote the matrix whose i, j-element is

$$\frac{\partial f_i}{\partial x_j}$$

For the function $F(t,y)$

$$F_t(t,y), \quad F_y(t,y)$$

will denote the derivatives with respect to t and y respectively, the fist being an n- vector and the second an n by n matrix. Since $F(t,y)$ is the solution of the differential equation, we have

$$F_t(t,y) = f(F(t,y))$$

Assuming that F is continuously differentiable with respect to y, we get

$$F_{t,y}(t,y) = f_x(F(t,y))F_y(t,y)$$

Since obviously

$$F_y(0,y) = I$$

we see that the derivative F_y *if* it exists must be a solution of the linear differential equation

$$\begin{array}{rcl} \frac{\partial}{\partial t}\psi(t,y) & = & f_x(F(t,y))\psi(t,y) \\ \psi(0,y) & = & I \end{array} \tag{2.15}$$

called the *variational equation.* In fact we have

Theorem 2.4 Let f in (2.11) be continuously differentiable. Then $F(t,y)$ is continuously differentiable with respect to y with the derivative satisfying (2.15).

Proof. Define

$$\theta(t,h) = F(t,y+h) - F(t,y)$$

We have

$$\theta(t,h) - \psi(t,h)h = \int_0^t \left(f(F(s,y+h)) - f(F(s,y))\right)\, ds -$$

$$- \int_0^t f_x(F(s,y))\psi(s,y)h\, ds = \int_0^t f_x(F(s,y))(\theta(s,h) - \psi(s,y)h)\, ds +$$

$$+ \int_0^t \left(\, f(F(s,y+h)) - f(F(s,y)) - f_x(F(s,y))\theta(s,h)\,\right)\, ds$$

Take an arbitrary $\epsilon > 0$. Since$f$ is differentiable there exists a $\delta > 0$ such that, for $|h| < \delta$, the last integral is bounded by

$$\int_0^t \epsilon |F(s,y+h) - F(s,y)| ds \leq C\epsilon |h|$$

for some constant C. Consequently we get the estimate

$$|\theta(t,h) - \psi(t,h)h| \leq \int_0^t f_x(F(s,y))|\theta(s,h) - \psi(s,h)h|\, ds + C\epsilon|h|$$

Gronwall's lemma then gives

$$|\theta(t,h) - \psi(t,h)h| \leq \tilde{C}\epsilon|h|$$

for a new constant $\tilde{C}$. From the definition of differentiability it then follows that

$$F_y(t,y) = \psi(t,y)$$

■

2.3 Series expansions

Consider again the differential equation

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0 \tag{2.16}$$

The solution $x(t)$ is continuously differentiable as a function of t, since it satisfies (2.16). Let us assume that f is continuously differentiable. Then the right hand side is a continuously differentiable function of t, which means that the left hand side is, which means that $x(t)$ is in fact twice continuously differentiable with respect to time. Differentiating using the chain rule gives

$$\ddot{x}(t) = f_x(x(t))\dot{x}(t) = f_x(x(t))f(x(t))$$

Defining the function

$$f^{(1)}(x) = f_x(x)f(x)$$

we can write

$$\ddot{x}(t) = f^{(1)}(x(t))$$

Let us now assume that f is twice continuously differentiable. Then the function $f^{(1)}(x)$ is once continuously differentiable. It follows that $\ddot{x}(t)$ is continuously differentiable with

$$x^{(3)}(t) = f^{(2)}(x(t))$$

where

$$f^{(2)}(x) = f_x^{(1)}(x)f(x)$$

Continuing in this fashion we have in fact proved

Theorem 2.5 Let f in (2.16) be k times continuously differentiable. Then the solution $x(t)$ is $k+1$ times continuously differentiable with

$$x^{(j+1)}(t) = f^{(j)}(x(t)), \quad j = 0, \ldots, k \tag{2.17}$$

where $f^{(j)}$ is defined recursively by

$$f^{(j)}(x) = f_x^{(j-1)}(x)f(x), \quad j = 1, \ldots, k, \quad f^{(0)}(x) = f(x) \tag{2.18}$$

Corollary 2.1 Let f in (2.16) be k times continuously differentiable. Then the solution $x(t)$ is given by

$$x(t) = x_0 + tf(x_0) + \frac{t^2}{2}f^{(1)}(x_0) + \cdots + \frac{t^k}{k!}f^{(k-1)}(x_0) + \frac{t^{k+1}}{(k+1)!}f^{(k)}(x(\xi)) \tag{2.19}$$

where $0 < \xi < t$.

Proof. Follows from a Taylor expansion. ∎

Example 2.1 Consider the pendulum equation

$$\begin{array}{rcl} \dot{x}_1 & = & x_2 \\ \dot{x}_2 & = & -\sin x_1 \end{array}, \quad x_0 = \begin{pmatrix} 0 \\ a \end{pmatrix}$$

We get

$$f^{(0)}(x) = f(x) = \begin{pmatrix} x_2 \\ -\sin x_1 \end{pmatrix}$$

$$f^{(1)}(x) = \begin{pmatrix} 0 & 1 \\ -\cos x_1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ -\sin x_1 \end{pmatrix} = \begin{pmatrix} -\sin x_1 \\ -x_2 \cos x_1 \end{pmatrix}$$

$$f^{(2)}(x) = \begin{pmatrix} -\cos x_1 & 0 \\ x_2 \sin x_1 & -\cos x_1 \end{pmatrix} \begin{pmatrix} x_2 \\ -\sin x_1 \end{pmatrix} = \begin{pmatrix} -x_2 \cos x_1 \\ x_2^2 \sin x_1 + \cos x_1 \sin x_1 \end{pmatrix}$$

This gives

$$f(x_0) = \begin{pmatrix} a \\ 0 \end{pmatrix}, \quad f^{(1)}(x_0) = \begin{pmatrix} 0 \\ -a \end{pmatrix}, \quad f^{(2)}(x_0) = \begin{pmatrix} -a \\ 0 \end{pmatrix}$$

with the series representation

$$x(t) = \begin{pmatrix} at - at^3/6 \\ a - at^2 \end{pmatrix} + O(t^4)$$

■

If f is in fact analytic, then one can show the stronger result

Theorem 2.6 Let f in (2.16) be analytic. Then x is an analytic function of t, given by

$$x(t) = x_0 + \sum_{k=1}^{\infty} \frac{t^k}{k!} f^{(k-1)}(x_0) \tag{2.20}$$

in a neighborhood of $t = 0$.

Example 2.2 Consider the scalar example

$$\dot{x} = x^2, \quad x(0) = x_0$$

In this case we get

$$f^{(n)}(x) = (n+1)! x^{n+2}$$

with the solution

$$x(t) = x_0 + \sum_{n=1}^{\infty} t^n x_0^{n+1} = \frac{x_0}{1 - x_0 t}$$

■

2.4 Exercises.

2.1 Show that the Riccati equation

$$\frac{d}{dt} P = AP + PA^T + Q - PC^T R^{-1} CP, \quad P(0) = P_0$$

has a solution on the time interval $[0, \epsilon]$ if ϵ is small enough.

2.2 Consider the scalar Riccati equation

$$\dot{p} = 1 + p^2, \quad p(0) = 0$$

Does it have a global solution ?

2.3 Use Theorem 2.2 to show that the scalar Riccati equation

$$\dot{p} = 2ap + q - p^2/r, \quad p(0) = p_0$$

has a solution on $[0, t_1]$ for any t_1, if $q \geq 0, \quad r > 0$.

2.4 Consider the differential equation

$$\frac{d}{dt}x = x^2, \quad x(0) = 1$$

Suppose the initial condition is changed to $x(0) = 1 + \epsilon$. Compute the change in $x(0.999)$ to first order in ϵ, using

a. the variational equation.

b. the exact solution of the differential equation for $x(0) = 1 + \epsilon$.

2.5 Consider the differential equation

$$\frac{d}{dt}x = 1 - x^2, \quad x(0) = 0$$

Suppose the initial condition is changed to $x(0) = \epsilon$. Compute the change in $x(1000)$ to first order in ϵ, using

a. the variational equation.

b. the exact solution of the differential equation for $x(0) = \epsilon$.

2.6 Consider the differential equation

$$\dot{x} = f(x, p), \quad x(0) = x_0$$

where f is differentiable. Prove that the solution $x(t, p)$ is differentiable with respect to the parameter p and compute an expression for the derivative.

Hint: Rewrite the problem so that p becomes an initial state, by introducing extra state variables.

Chapter 3

Canonical forms and exact linearization.

In linear system theory it is well known that controller design is particularly easy when the system is described by a controller canonical form. It is then natural to ask if a similar canonical form can be achieved for a nonlinear system. We will approach this question by first looking at the input-output properties. This will lead to the definition of the relative degree of a nonlinear system.

3.1 The relative degree of a nonlinear system

Consider a nonlinear system of the form

$$\begin{aligned} \dot{x} &= f(x) + g(x)u && (3.1) \\ y &= h(x) && (3.2) \end{aligned}$$

where x is an n-vector, u and y are m-vectors, and f, g and h are infinitely differentiable functions. Here g is an $n \times m$ matrix of differentiable functions and h an m-vector of differentiable functions. Let g_1,..,g_m denote the column vectors of g.

An important structural question for such a system is the extent to which inputs directly affect outputs or their derivatives. Since we have assumed that the right hand side of (3.2) does not depend on u, the output is not directly dependent on any of the input signals. Differentiating the i:th output we get, using the chain rule

$$\dot{y}_i = h_{i,x}(x)(f(x) + g(x)u) \qquad (3.3)$$

where $h_{i,x}$ is the row vector with elements

$$h_{i,x}(x) = \left(\frac{\partial h_i(x)}{\partial x_1}, \ldots, \frac{\partial h_i(x)}{\partial x_n} \right)$$

Let us introduce the operators

$$L_f = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i}, \quad L_{g_k} = \sum_{i=1}^{n} g_{ik} \frac{\partial}{\partial x_i} \tag{3.4}$$

We refer to L_f as the *Lie derivative* in the direction f. We can then rewrite (3.3) as

$$\dot{y}_i = L_f h_i + \sum_{k=1}^{m} u_k \, L_{g_k} h_i \tag{3.5}$$

Before proceeding further let us consider some of the properties of the Lie derivative. As shown by (3.4) it is a first order derivative operator. Applying one Lie derivative to another gives an operator involving second order derivatives:

$$L_f L_g = \sum_i f_i \frac{\partial}{\partial x_i} \sum_j g_j \frac{\partial}{\partial x_j} == \sum_{i,j} \left(f_i \frac{\partial g_j}{\partial x_i} \right) \frac{\partial}{\partial x_j} + \sum_{i,j} f_i g_j \frac{\partial^2}{\partial x_i \partial x_j}$$

However, if we take $L_f L_g - L_g L_f$, then the second order derivatives will cancel out, and we get a new first order operator, that can be interpreted as a Lie derivative.

$$L_f L_g - L_g L_f = \sum_j \left(\sum_i f_i \frac{\partial g_j}{\partial x_i} - g_i \frac{\partial f_j}{\partial x_i} \right) \frac{\partial}{\partial x_j}$$

The expression within parenthesis is called the *Lie bracket* of f and g and is denoted $[f, g]$:

$$[f, g] = g_x f - f_x g = \sum_i f_i \frac{\partial g_j}{\partial x_i} - g_i \frac{\partial f_j}{\partial x_i}$$

We have thus shown the following formula

$$L_f L_g - L_g L_f = L_{[f,g]} \tag{3.6}$$

Sometimes it is useful to consider repeated Lie brackets of the form

$$[f, g], \;\; [f, [f, g]], \;\; [f, [f, [f, g]]], \ldots$$

The j times repeated Lie bracket of this form is denoted $(ad^j f, g)$, for instance:

$$(ad^3 f, g) = [f, [f, [f, g]]]$$

After this parenthesis about Lie derivatives we go back to (3.5). If $L_{g_k} h_i$ is not identically zero, then obviously the control u_k will directly influence $\dot{y}_i$, at lest for some x. Suppose $L_{g_k} h_i = 0$ for all k and all x. Then we have $\dot{y}_i = L_f h_i$ and we can continue differentiating.

$$\ddot{y}_i = L_f^2 h_i + \sum_{k=1}^{m} u_k \, L_{g_k} L_f h_i$$

As before we can see that u_k influences $\ddot{y}_i$ directly if $L_{g_k} L_f h_i$ is not identically zero. If $L_{g_k} L_f h_i = 0$ for all k and x, we continue the differentiations. There are two possibilities: either no time derivative of y_i is directly dependent on the output, or else there is a smallest integer ν_i such that $y_i^{(\nu_i)}$ depends directly on some u_k. In the latter case we have

$$\begin{aligned} L_{g_k}L_f^j h_i &\equiv 0, \quad k=1,\dots,m, \quad j=0,\dots,\nu_i-2 \\ L_{g_k}L_f^{\nu_i-1} h_i &\not\equiv 0, \quad \text{some } k,\ 1 \le k \le m \end{aligned} \tag{3.7}$$

Suppose that we have done the above procedure for all output signals and obtained numbers $\nu_1,\dots,\nu_m$, satisfying (3.7). We can then write the input-output relationship as

$$\begin{pmatrix} y_1^{(\nu_1)} \\ \vdots \\ y_m^{(\nu_m)} \end{pmatrix} = d(x) + R(x)u \tag{3.8}$$

where d and R are given by

$$d(x) = \begin{pmatrix} L_f^{\nu_1} h(x) \\ \vdots \\ L_f^{\nu_m} h(x) \end{pmatrix}, \quad R(x) = \begin{pmatrix} L_{g_1}L_f^{\nu_1-1}h_1 & \dots & L_{g_m}L_f^{\nu_1-1}h_1 \\ \vdots & & \vdots \\ L_{g_1}L_f^{\nu_m-1}h_m & \dots & L_{g_m}L_f^{\nu_m-1}h_m \end{pmatrix} \tag{3.9}$$

From the definition of the ν_i it follows that each row of R has at least one element which is not identically zero. If R is nonsingular calculations become especially simple as we shall see below.

Definition 3.1 We say that the system (3.1,3.2) has vector *relative degree* $(\nu_1,\dots,\nu_m)$ at x_0 if (3.7) is satisfied and $R(x_0)$ is nonsingular

The main point in all the calculations we have made lies in the following fact. Suppose the system has a relative degree. Then we can use the state feedback

$$u = R(x)^{-1}(\bar{u} - d(x)) \tag{3.10}$$

where we regard $\bar{u}$ as a new input signal. The resulting dynamics is

$$\begin{pmatrix} y_1^{(\nu_1)} \\ \vdots \\ y_m^{(\nu_m)} \end{pmatrix} = \bar{u} \tag{3.11}$$

This is a decoupled linear relation from the new input to the output. Using linear design techniques it is now possible to get any dynamics from input to output. Our calculations can be summarized in the following theorem:

Theorem 3.1 Input-output linearization. A system having relative degree can be given linear dynamics from input to output by using state feedback.

Proof. Follows directly from (3.8) and (3.10). ∎

Example 3.1 Consider the rigid body of Example 1.4 with external moments along only two of the principal axes. We consider these moments to be input signals.

$$\begin{aligned}\dot{x}_1 &= a_1 x_2 x_3 &(3.12)\\ \dot{x}_2 &= a_2 x_1 x_3 + u_1 &(3.13)\\ \dot{x}_3 &= a_3 x_1 x_2 + u_2 &(3.14)\end{aligned}$$

Assume that the outputs are the first and second angular velocities:

$$y_1 = x_1, \quad y_2 = x_2 \tag{3.15}$$

Since

$$L_{g_1} = \frac{\partial}{\partial x_2}, \quad L_{g_2} = \frac{\partial}{\partial x_3}$$

$$L_f = a_1 x_2 x_3 \frac{\partial}{\partial x_1} + a_2 x_1 x_3 \frac{\partial}{\partial x_2} + a_3 x_1 x_2 \frac{\partial}{\partial x_3}$$

we have

$$L_{g_1} h_1 = L_{g_1} x_1 = 0, \quad L_{g_2} h_1 = L_{g_2} x_1 = 0$$

and

$$L_{g_1} L_f h_1 = L_{g_1} a_1 x_2 x_3 = a_1 x_3, \quad L_{g_2} L_f h_1 = L_{g_2} a_1 x_2 x_3 = a_1 x_2$$

showing that $\nu_1 = 2$ if $a_1 \neq 0$. Also we have

$$L_{g_1} h_2 = L_{g_1} x_2 = 1, \quad L_{g_2} h_2 = L_{g_2} x_2 = 0$$

so that $\nu_2 = 1$ and

$$R(x) = \begin{pmatrix} a_1 x_3 & a_1 x_2 \\ 1 & 0 \end{pmatrix}$$

If $a_1 \neq 0$, i.e. if the moments of inertia around the second and third axes are not equal, then we see that the system has vector relative degree $(2, 1)$ if $x_2 \neq 0$. If on the other hand $a_1 = 0$, then all time derivatives of y_1 are zero, so there is no relative degree. ■

A major problem with the input-output linearization comes from the fact that the linear input-output relationship (3.11) might not show all the dynamics of the closed loop system. This is the case if the order of the system (3.11), i.e. $\nu_1 + \cdots + \nu_m$ is less than n, the number of state variables. There must then be some dynamics which is unobservable from the output. If this hidden dynamics is unstable it will not be possible to use the input-output linearization. To investigate this question it is necessary to transform the system into a canonical form with all state variables present.

3.2 A canonical form

Our discussion of relative degree and the example of the previous section suggest that it might be natural to regard the outputs and their derivatives as state variables. Therefore we define new variables in the following way.

$$\begin{array}{rcl} \zeta & = & (y_1, \dot{y}_1, \ldots, y_1^{(\nu_1-2)}, y_2, \dot{y}_2, \ldots, y_2^{(\nu_2-2)}, \ldots, y_m, \dot{y}_m, \ldots, y_m^{(\nu_m-2)})^T \\ \xi & = & (y_1^{(\nu_1-1)}, \ldots, y_m^{(\nu_m-1)})^T \\ \eta & = & q(x) \end{array} \tag{3.16}$$

or equivalently

$$\begin{array}{rcl} \zeta & = & (h_1, L_f h_1, \ldots, L_f^{(\nu_1-2)} h_1, \ldots, h_m, L_f h_m, \ldots, L_f^{(\nu_m-2)} h_m)^T \\ \xi & = & (L_f^{(\nu_1-1)} h_1, \ldots, L_f^{(\nu_m-1)} h_m)^T \\ \eta & = & q(x) \end{array} \tag{3.17}$$

where q is some function which is unspecified so far. If the relation between these variables and x is invertible, then the new variables can be used as state variables. We then get an interesting canonical form for the system.

Theorem 3.2 If the system (3.1), (3.2) has relative degree $(\nu_1, \ldots, \nu_m)$ at x_0, then $q(x)$ in the variable change (3.17) can be chosen so that (3.17) is invertible in a neighborhood of x_0. In the new variables the state space description is

$$\begin{array}{rcl} \dot{\zeta} & = & M\zeta + N\xi \\ \dot{\xi} & = & \psi_1(\zeta, \xi, \eta) + R(\phi(\zeta, \xi, \eta))u \\ \dot{\eta} & = & \psi_2((\zeta, \xi, \eta) + \psi_3(\zeta, \xi, \eta)u \\ y & = & H_1\zeta + H_2\xi \end{array} \tag{3.18}$$

where M is a block diagonal matrix, where the i:th block has dimension $(\nu_i - 1) \times (\nu_i - 1)$:

$$M = \begin{pmatrix} M_{11} & 0 & \ldots & 0 \\ 0 & M_{22} & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & & & M_{mm} \end{pmatrix}, \quad \text{with } M_{ii} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

N is also a block diagonal matrix with m blocks. Each block has $\nu_i - 1$ elements and is of the form

$$N_{ii} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

The matrices H_1 and H_2 are implicitly defined by (3.17). The function ϕ is the inverse of (3.17)

To prove this theorem it is necessary to do some preliminary work. Introduce the matrices

$$S(x) = \Big(h_{1,x}, \ldots, (L_f^{\nu_1-1} h_1)_x, \ldots, h_{m,x}, \ldots, (L_f^{\nu_m-1} h_m)_x\Big)^T \tag{3.19}$$

$$T(x) = \big(g_1, \ldots, g_m, [f, g_1], \ldots, [f, g_m], \ldots, (ad^{r-1} f, g_1), \ldots, (ad^{r-1} f, g_m)\big) \tag{3.20}$$

where $r = \max(\nu_1, \ldots, \nu_m)$. We have the following fundamental fact

Proposition 3.1 If the system (3.1), (3.2) has relative degree $(\nu_1, \ldots, \nu_m)$ at x_0, then the matrix $S(x_0)T(x_0)$ has rank $\nu_1 + \cdots + \nu_m$ (= number of rows).

Proof. Consider a case where $m = 2$, $\nu_1 = 2$, $\nu_2 = 1$. Then

$$ST = \begin{pmatrix} h_{1,x} \\ (L_f h_1)_x \\ h_{2,x} \end{pmatrix} (g_1, g_2, [f, g_1], [f, g_2]) =$$

$$\begin{pmatrix} L_{g_1} h_1 & L_{g_2} h_1 & L_{[f,g_1]} h_1 & L_{[f,g_2]} h_1 \\ L_{g_1} L_f h_1 & L_{g_2} L_f h_1 & * & * \\ L_{g_1} h_2 & L_{g_2} h_2 & * & * \end{pmatrix}$$

where $*$ denotes elements whose values are unimportant. Now, since $\nu_1 = 2$, we have $L_{g_1} h_1 = 0$, $L_{g_2} h_1 = 0$. Using (3.6) we have

$$L_{[f,g_1]} h_1 = L_f L_{g_1} h_1 - L_{g_1} L_f h_1 = -L_{g_1} L_f h_1$$

$$L_{[f,g_2]} h_1 = L_f L_{g_2} h_1 - L_{g_2} L_f h_1 = -L_{g_2} L_f h_1$$

so that

$$S(x_0)T(x_0) = \begin{pmatrix} 0 & 0 & -L_{g_1} L_f h_1 & -L_{g_2} L_f h_1 \\ L_{g_1} L_f h_1 & L_{g_2} L_f h_1 & * & * \\ L_{g_1} h_2 & L_{g_2} h_2 & * & * \end{pmatrix}$$

The 2×2 block matrix to the lower left is exactly $R(x_0)$ and is therefore nonsingular. The two elements to the upper right constitute the first row of $R(x_0)$, so they can not both be zero. Consequently $S(x_0)T(x_0)$ has full rank. This case ($m = 2$, $\nu_1 = 2$, $\nu_2 = 1$) can easily be extended to the general one. By repeated use of (3.6) and possibly permutation of the columns, the matrix can be brought into block triangular form, where the blocks on the diagonal consist of rows of $R(x_0)$. The matrix must then have full rank. ∎

From this proposition we can draw the following conclusions.

Proposition 3.2 If the system (3.1), (3.2) has relative degree $(\nu_1, \ldots, \nu_m)$ at x_0, then

1. $\nu_1 + \cdots + \nu_m \leq n$
2. $S(x_0)$ has linearly independent rows
3. there are at least $\nu_1 + \cdots + \nu_m$ independent columns in $T(x_0)$

Proof. From Proposition 3.1 and the properties of the matrix rank, it follows that

$$\nu_1 + \cdots + \nu_m = \operatorname{rank} S(x_0)T(x_0) \leq \min(\operatorname{rank} S(x_0), \operatorname{rank} T(x_0))$$

All three statements follow immediately. ∎

We can now prove the main result.

Proof.(of Theorem 3.2) The Jacobian (at x_0) of the coordinate change (3.17) is (after some reordering of the rows)

$$\begin{pmatrix} S(x_0) \\ q_x(x_0) \end{pmatrix}$$

We know from Proposition 3.2 that $S(x_0)$ has linearly independent rows. Since we are free to choose q, we can choose it so that the rows of q_x are linearly independent of those of S. The Jacobian is then nonsingular and it follows, from the implicit function theorem, that the variable change (3.17) is invertible close to x_0. ■

The canonical form (3.18) suggests an interesting choice of controller. Suppose v is an m-vector of reference signals for the output y. Then the control law

$$u = k(\zeta,\xi,\eta,v) = R(\phi(\zeta,\xi,\eta))^{-1}(-\psi_1(\zeta,\xi,\eta) + \Xi_1(\zeta,\xi) + \Xi_2(\zeta,\xi)v) \quad (3.21)$$

gives the dynamics

$$\begin{array}{rcl} \dot{\zeta} & = & M\zeta + N\xi \\ \dot{\xi} & = & \Xi_1(\zeta,\xi) + \Xi_2(\zeta,\xi)v \\ y & = & H\zeta \end{array} \quad (3.22)$$

from reference signal to output. Notice that the η variable is not included in this dynamics. The functions Ξ_1 and Ξ_2 can be chosen arbitrarily. There are several interesting possibilities.

Noninteracting control

If the functions Ξ_1 and Ξ_2 in the control law (3.21) are chosen so that the i:th component is of the form

$$(\Xi_1(\zeta,\xi))_i = a_i(y_i, \dot{y}_i, \ldots, y_i^{(\nu_i-1)})$$

$$(\Xi_2(\zeta,\xi)v)i = b_i(y_i, \dot{y}_i, \ldots, y_i^{(\nu_i-1)})\, v_i$$

then the relation between v and y breaks down into m relations of the form

$$y_i^{(\nu_i)} = a_i(y_i, \dot{y}_i, \ldots, y_i^{(\nu_i-1)}) + b_i(y_i, \dot{y}_i, \ldots, y_i^{(\nu_i-1)})\, v_i$$

We have then achieved *noninteracting control.* The i:th reference signal affects only the i:th output.

Linear reference-output relation

Let the functions Ξ_1 and Ξ_2 in the control law (3.21) be chosen in the following way.

$$\Xi_1(\zeta,\xi) = \tilde{M}\zeta + \tilde{N}\xi$$

$$\Xi_2(\zeta,\xi) = \tilde{R}$$

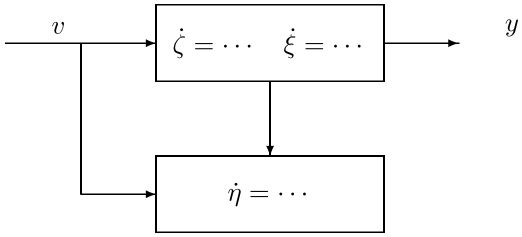


Figure 3.1: Structure of the system

where $\tilde{M}$, $\tilde{N}$ and $\tilde{R}$ are constant matrices of appropriate dimensions. The resulting dynamics is then

$$\begin{array}{rcl} \dot{\zeta} & = & M\zeta + N\xi \\ \dot{\xi} & = & \tilde{M}\zeta + \tilde{N}\xi + \tilde{R}v \\ y & = & H\zeta \end{array} \tag{3.23}$$

which is linear from v to y.

3.3 The zero dynamics

The controllers discussed in the previous section have a peculiar property. If we consider equation (3.22), we see that the variable η of (3.18) does not effect the reference–output relationship. Writing down all of the dynamics gives

$$\begin{array}{rcl} \dot{\zeta} & = & M\zeta + N\xi \\ \dot{\xi} & = & \Xi_1(\zeta,\xi) + \Xi_2(\zeta,\xi)v \\ \dot{\eta} & = & \psi_2((\zeta,\xi,\eta) + \psi_3(\zeta,\xi,\eta)k(\zeta,\xi,\eta,v) \\ y & = & H\zeta \end{array} \tag{3.24}$$

Graphically the situation described by (3.24) is shown in figure 3.1. We see that the η dynamics acts as a "sink". It does not affect the output put is (possibly) affected by v, ζ and ξ. This is something that we recognize from linear system theory – part of the dynamics is unobservable.

Let us consider the situation where $v = 0$, so that the desired output is $y \equiv 0$. It is then natural to choose Ξ_1 so that

$$\Xi_1(0,0) = 0$$

Suppose the system is initialized with

$$y_i(0) = 0,\ \dot{y}_i(0) = 0, \ldots, y_i^{(\nu_i - 1)}(0) = 0, \quad i = 1, \ldots, m$$

We will then have

$$\zeta \equiv 0, \quad \xi \equiv 0$$

$$\dot{\eta} = \psi_2(0,0,\eta) + \psi_3(0,0,\eta)k(0,0,\eta,0) \tag{3.25}$$

Definition 3.2 The dynamics described by (3.25) is called the *zero dynamics* of the system.

The term “zero dynamics” is motivated by the linear case.

Example 3.2 Consider the system (observer canonical form)

$$\dot{x} = \begin{pmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ b_2 \\ b_3 \end{pmatrix} u$$

$$y = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$$

with transfer function

$$G(s) = \frac{b_2 s + b_3}{s^3 + a_1 s^2 + a_2 s + a_3}$$

We assume that $b_2 \neq 0$. Then $\dot{y}$ does not depend directly on u, but $\ddot{y}$ does, so the relative degree is $\nu_1 = 2$. Introducing

$$\zeta_1 = x_1, \quad \xi_1 = -a_1 x_1 + x_2, \quad \eta = x_3$$

we have the canonical form

$$\begin{aligned} \dot{\zeta}_1 &= \xi_1 \\ \dot{\xi}_1 &= -a_1\xi_1 - a_2\zeta_1 + \eta_3 + b_2 u \\ \dot{\eta} &= -a_3\zeta_1 + b_3 u \end{aligned}$$

The zero dynamics corresponds to the control that keeps ζ_1 and ξ_1 identically zero, i. e. $u = -\eta_3/b_2$. The zero dynamics is then

$$\dot{\eta}_3 = -\frac{b_3}{b_2}\eta_3$$

The eigenvalue of the zero dynamics $(-b_3/b_2)$ is thus equal to the zero of the transfer function. ■

This example is easily extended to single-input-single-output linear systems of arbitrary order and relative degree. The conclusion is the same: the eigenvalues of the zero dynamics are the zeroes of the transfer function. By analogy with the linear case the term zero dynamics is used also in the nonlinear case. The linear analogy also explains the following definition.

Definition 3.3 A nonlinear system is called *minimum phase* if the origin is an asymptotically stable point of the zero dynamics.

To see the significance of the minimum phase property, we consider another example.

Example 3.3 Consider the system

$$\begin{array}{rcl} \dot{x}_1 & = & -x_2^2 + u \\ \dot{x}_2 & = & u \\ y & = & x_1 \end{array} \tag{3.26}$$

This system already has the canonical form. To keep y identically zero, one has to use the control law

$$u = x_2^2$$

The zero dynamics is then

$$\dot{x}_2 = x_2^2$$

which is clearly unstable, so the system is non minimum phase. ■

This example clearly shows that control of non minimum phase systems can be tricky. A control that keeps the output identically zero is unacceptable since it makes certain internal states (and the control signal) grow without bound.

The problems with the zero dynamics dissappear if the relative degree satisfies

$$\nu_1 + \cdots + \nu_m = n$$

In this case the η vector has dimension zero, so there is no zero dynamics. Our earlier example 3.1 is such a case.

Example 3.4 Consider again example 3.1. The coordinate change

$$\zeta_1 = x_1, \quad \xi_1 = a_1 x_2 x_3, \quad \xi_2 = x_2$$

gives the dynamics

$$\begin{array}{rcl} \dot{\zeta}_1 & = & \xi_1 \\ \dot{\xi}_1 & = & \frac{a_2 \zeta_1 \xi_1^2}{a_1 \xi_2^2} + \frac{\xi_1}{\xi_2} u_1 + a_1 a_3 \zeta_1 \xi_2^2 + a_1 \xi_2 u_2 \\ \dot{\xi}_2 & = & \frac{a_2 \zeta_1 \xi_1}{a_1 \xi_2} + u_1 \end{array}$$

The feedback

$$u_1 = -\frac{a_2 \zeta_1 \xi_1}{a_1 \xi_2} + v_2$$

$$u_2 = \frac{1}{a_1 \xi_2} \left(\frac{\xi_1}{\xi_2} v_2 - a_1 a_3 \zeta_1 \xi_2^2 + v_1 \right)$$

gives the decoupled closed loop dynamics

$$\ddot{y}_1 = v_1, \quad \dot{y}_2 = v_2$$

Further linear feedback can then give arbitrary closed loop poles. There is no hidden dynamics. ■

3.4 Controller canonical form.

The last section showed us that there are great advantages with systems where the relative degree satisfies

$$\nu_1 + \cdots + \nu_m = n \tag{3.27}$$

so that there is no zero dynamics. Suppose one starts with a system

$$\dot{x} = f(x) + g(x)u \tag{3.28}$$

Is it then possible to *choose* an output so that the relative degree satisfies (3.27)? To answer that question we will specialize to the case $m = 1$, i.e. the single-input-single-output case.

From (3.7) it follows that we require a function $h(x)$ satisfying

$$\begin{array}{rcl} L_g L_f^j h & \equiv & 0, \quad j = 0, \ldots, n-2 \\ L_g L_f^{n-1} h & \not\equiv & 0 \end{array} \tag{3.29}$$

This can be rewritten using (3.6):

$$L_{[f,g]}h = L_f L_g h - L_g L_f h = 0$$

$$L_{[f,[f,g]]}h = L_f L_{[f,g]}h - L_{[f,g]}L_f h = -L_f L_g L_f h + L_g L_f^2 h = 0$$

and so on. We find that (3.29) is equivalent to

$$L_{(ad^k f,g)}h = 0, \quad k = 0, 1, \ldots, n-2 \tag{3.30}$$

$$L_{(ad^{n-1} f,g)}h \neq 0 \tag{3.31}$$

This is a system of partial differential equations that has to be solved. It is helpful to look at its geometric interpretation. For $n = 3$ we have the conditions

$$L_g h = 0, \quad L_{[f,g]}h = 0$$

They are satisfied if we can find a family of surfaces of the form $h(x) = c$ such that g and $[f, g]$ are tangent to the surfaces at every point, see figure 3.2. It turns out that this construction can only be carried out if the vectors satisfy certain conditions. To discuss this we need the following definition.

Definition 3.4 A collection of vector fields $f_1(x), \ldots, f_p(x)$ is called *involutive* if all the Lie brackets are linear combinations of the f_i, i. e.

$$[f_i, f_j](x) = \sum_{k=1}^{p} \gamma_k(x) f_k(x), \quad i, j = 1, \ldots, p$$

where the γ_k are infinitely differentiable scalar functions.

We also need the following fact about Lie brackets.

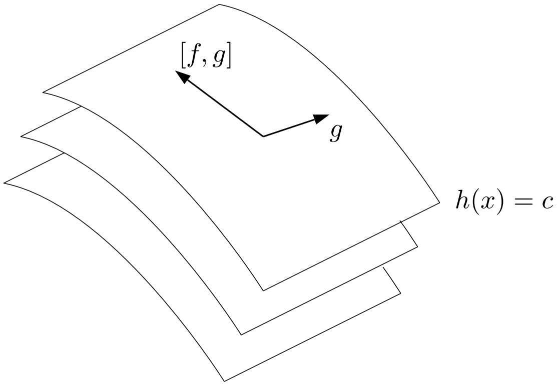


Figure 3.2: Geometric interpretation of the conditions $L_g h = 0$, $L_{[f,g]}h = 0$.

Proposition 3.3 Let the solution of

$$\dot{x} = f(x), \quad x(0) = x_0$$

after t units of time be x_a, let the solution of

$$\dot{x} = g(x), \quad x(0) = x_a$$

at $t = h$ be x_b, let the solution of

$$\dot{x} = -f(x), \quad x(0) = x_b$$

at $t = h$ be x_c and let finally the solution of

$$\dot{x} = -g(x), \quad x(0) = x_c$$

at $t = h$ be x_d (see Figure 3.3). Then

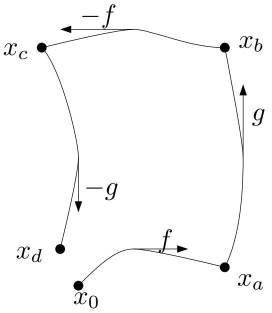


Figure 3.3: Geometric construction to interpret the Lie bracket.

$$x_d = x_0 + h^2[f,g](x_0) + O(h^3)$$

Proof. A Taylor expansion gives

$$x_a = x_0 + h\dot{x}(0) + \frac{h^2}{2}\ddot{x}(0) + O(h^3)$$

Since $\dot{x} = f$ and $\ddot{x} = f_x\dot{x} = f_x f$ this can be written

$$x_a = x_0 + hf(x_0) + \frac{h^2}{2}f_x(x_0)f(x_0) + O(h^3)$$

Analogous calculations give

$$\begin{aligned}
x_b &= x_a + hg(x_a) + \frac{h^2}{2}g_x(x_a)g(x_a) + O(h^3)\\
x_c &= x_b - hf(x_b) + \frac{h^2}{2}f_x(x_b)f(x_b) + O(h^3)\\
x_d &= x_c - hg(x_c) + \frac{h^2}{2}g_x(x_c)g(x_c) + O(h^3)
\end{aligned}$$

The right hand sides of these expressions can be evaluated at x_0. For instance

$$g(x_a) = g(x_0) + g_x(x_0)(x_a - x_0) + O(h^2) = g(x_0) + hg_x(x_0)f(x_0) + O(h^2)$$

Performing analogous calculations for x_c and x_d gives

$$\begin{aligned}
x_a &= x_0 + hf + \frac{h^2}{2}f_x f + O(h^3)\\
x_b &= x_a + hg + h^2 g_x f + \frac{h^2}{2}g_x g + O(h^3)\\
x_c &= x_b - hf - h^2 f_x(f+g) + \frac{h^2}{2}f_x f + O(h^3)\\
x_d &= x_c - hg - h^2 g_x(f+g-f) + \frac{h^2}{2}g_x g + O(h^3)
\end{aligned}$$

where the right hand side is evaluated at x_0. Substituting each equation into the one below finally gives

$$x_d = x_0 + h^2(g_x f - f_x g) + O(h^3) = x_0 + h^2[f,g] + O(h^3$$

■

This result has immediate consequences for the construction of Figure 3.2. Consider the construction of Proposition 3.3 but with f and g replaced by g and $[f,g]$. If g and $[f,g]$ are both tangent to $h(x) = c$ at every point, then all solutions to $\dot{x} = g$ and $\dot{x} = [f,g]$ lie in $h(x) = c$. From Proposition 3.3 the Lie bracket $[g,[f,g]]$ must then be tangent to $h(x) = c$. It will then be a linear combination of g and $[f,g]$ (with x-dependent coefficients). The geometric picture of Figure 3.2 thus implies that g and $[f,g]$ are involutive.

Theorem 3.3 The system (3.30) with the condition (3.31) has a solution in a neighborhood of a point x_0 if and only if

1. the vectors
$$(ad^k f, g)(x_0), \quad k = 0, \dots, n-1 \tag{3.32}$$
are linearly independent

2. the vector fields
$$(ad^k f, g), \quad k = 0, \dots, n-2 \tag{3.33}$$
are involutive in a neighborhood of x_0.

Proof. (sketch) First we note that the necessity of condition 1 follows from the third statement of Proposition 3.2.

For the other part of the theorem it is helpful to consider again the geometric interpretation of figure 3.2. In the illustrated three-dimensional case g and $[f, g]$ have to be tangent to $h(x) = c$. For a general n the surface $h(x) = c$ has to be tangent to g, $[f, g]$,...,$(ad^{n-2} f, g)$ at all points.

Now consider what happens if a curve is generated which is tangent to g, then to $[f, g]$, then to $-g$ and finally to $-[f, g]$. Since all these vector fields are tangent to the surface $h(x) = c$, the curve will remain in the surface if it starts there. On the other hand it follows from Proposition 3.3 that the resulting movement will be in the direction $[g, [f, g]]$. Consequently $[g, [f, g]]$ also has to be tangent to $h(x) = c$. This argument can be extended to show that all brackets formed among g, $[f, g]$,...,$(ad^{n-2} f, g)$ have to be tangent, showing that these vector fields have to be involutive. We have thus sketched the necessity part of the theorem.

That our conditions are also sufficient follows from a celebrated theorem by Frobenius, forming one of the cornerstones of differential geometry. ■

After calculation of the function h we can introduce new state variables according to (3.17). In this special case we get
$$z_k = \left(L_f^{k-1} h\right)(x), \quad k = 1, \dots, n \tag{3.34}$$
with the state space description
$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_n &= L_f^n h + (-1)^{n-1} L_{(ad^{n-1} f, g)} h \; u \end{aligned} \tag{3.35}$$

We will call this form *controller canonical form* in analogy with the linear case. In our discussion of relative degree we assumed that the output is $y = z_1$. However, the controller canonical form can be useful no matter what the output is. In many cases we can regard $h(x)$ as a fictitious output, used as an intermediate step in the calculations.

3.4.1 Construction of the coordinate transformation.

We now turn to the actual computation of the function h whose existence is guaranteed by Theorem 3.3. The procedure is also illustrated in figure 3.4.

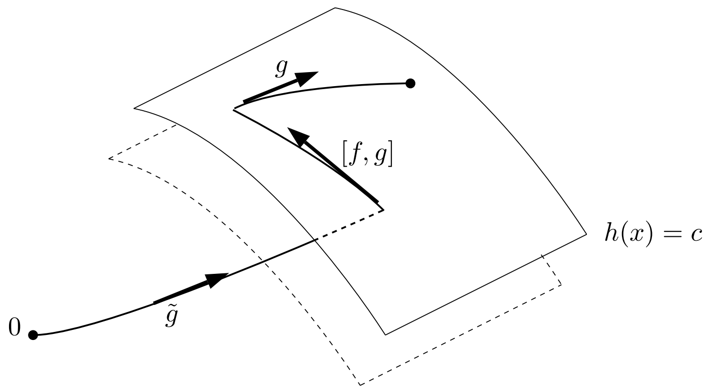


Figure 3.4: Construction of coordinate transformation.

Take an arbitrary vector field $\tilde{g}(x)$ which is linearly independent of the vectors in (3.33) . Let the solution of $\dot{x} = \tilde{g}(x)$ with initial value x_0 be denoted $\pi_1(t, x_0)$. Also let

$$\pi_2(t, x_0), \dots, \pi_n(t, x_0)$$

denote the solutions of

$$\dot{x} = (ad^{n-2}f, g), \dots, \dot{x} = [f.g], \dot{x} = g(x)$$

respectively. Consider the following procedure. Starting at $x = 0$, follow the solution of $\dot{x} = \tilde{g}$ for t_1 units of time to the point $x = \pi_1(t_1, 0)$. Then follow the solutions of $\dot{x} = (ad^{n-2}f, g)$, $\dot{x} = (ad^{n-3}f, g)$ etc. for $t_2, t_3, \dots$ units of time respectively. Finally after following the solution curve of $\dot{x} = g$ for t_n units of time, the point

$$x = \psi(t_1, \dots, t_n) \overset{\triangle}{=} \pi_n\Big(t_n, \pi_{n-1}\big(t_{n-1}, \dots \pi_1(t_1, 0)\big)\Big) \tag{3.36}$$

is arrived at. It is easy to see that the Jacobian of ψ evaluated at $t_i = 0$ is simply

$$\begin{pmatrix} \tilde{g} & (ad^{n-2}f, g) & \dots & [f, g] & g \end{pmatrix}$$

Since the columns are linearly independent, the Jacobian of ψ is nonsingular at the origin. It follows from the inverse function theorem that there is a neighborhood of the origin where ψ is invertible. Then it is possible to solve the equation $\psi(t_1, \dots, t_n) = x$ for t_1, giving

$$t_1 = h(x)$$

To show the procedure, a simple example is worked out.

Example 3.5 The following system is given.

$$\begin{aligned} \dot{x}_1 &= x_2 + x_2 u \\ \dot{x}_2 &= -x_2 + u \end{aligned}$$

The vectors of (3.32) are then

$$g = \begin{pmatrix} x_2 \\ 1 \end{pmatrix} \quad [f,g] = \begin{pmatrix} -x_2 - 1 \\ 1 \end{pmatrix}$$

which are linearly independent if $x_2 \neq -1/2$. Taking $\tilde{g} = (1 \quad 0)^T$ gives

$$\psi(t_1, t_2) = \begin{pmatrix} t_1 + t_2^2/2 \\ t_2 \end{pmatrix}$$

and solving for t_1 gives

$$h(x) = x_1 - x_2^2/2$$

From (3.34) the coordinate change is then

$$\begin{aligned} z_1 &= x_1 - x_2^2/2 \\ z_2 &= x_2 + x_2^2 \end{aligned}$$

or equivalently

$$\begin{aligned} x_1 &= 1/4 + z_1 + z_2/2 - \sqrt{z_2 + 1/4}\Big/2 \\ x_2 &= -1/2 + \sqrt{z_2 + 1/4} \end{aligned}$$

The canonical form shown by (3.35) is then

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= -1/2 - 2z_2 + \sqrt{z_2 + 1/4} + 2\sqrt{z_2 + 1/4}\ u \end{aligned}$$

Clearly the relationship between x_2 and z_2 is only invertible if $x_2 > -1/2$, so the transformation is not a global one. ■

3.5 Exact linearization.

We saw earlier (equation (3.23)) that it is possible to get a linear relationship between reference and output. For the special case we are discussing now we get the following result.

Theorem 3.4 Suppose the system (3.28) with a scalar input satisfies the conditions of Theorem 3.3. Then there exists a nonlinear transformation

$$z = T(x)$$

and a nonlinear feedback

$$u = k_1(x) + k_2(x)v$$

such that in the new variables the system has the form

$$\dot{z} = Az + Bv$$

i.e. it is a linear system.

Proof. First transform the system to controller canonical form (3.35) as described in the previous section. Then apply the nonlinear feedback

$$u = \frac{1}{\beta(x)}(-\alpha(x) + v)$$

(The resulting system has only pure integrators. This is sometimes called the Bronowsky canonical form. Of course it is possible to get any pole placement by adding suitable terms.) ■

3.6 Exact linearization – the multi-input case

The results for the single input case can be generalized to the case where u in

$$\dot{x} = f(x) + g(x)\,u \tag{3.37}$$

is an m-vector. One then has to consider

$$\begin{aligned} G_0 &= \operatorname{span}(g_1, \dots, g_m) \\ G_1 &= \operatorname{span}(g_1, \dots, g_m, [f, g_1], \dots, [f, g_m]) \\ &\vdots \\ G_{n-1} &= \operatorname{span}((\operatorname{ad}^k f, g_j),\ \ k = 0, \dots, n-1,\ \ j = 1, \dots, m) \end{aligned}$$

Theorem 3.5 Assume that $g(x_0)$ has rank m. For the system (3.37) it is possible to introduce an output $y = h(x)$ so that the relative degree satisfies $\nu_1 + \cdots + \nu_m = n$ in a neighborhood of x_0 if and only if

1. all G_i have constant rank near x_0
2. G_{n-1} has rank n.
3. $G_0, \dots, G_{n-2}$ are involutive.

We illustrate the theorem with an example.

Example 3.6 Consider the control of a missile, figure 3.5. The components of the velocity along the x-, y- and z-axes are u, v and w respectively. The angular velocities are p, q and r. The forward velocity usually changes fairly slowly and is therefore assumed to be constant. The force equations along the y- and z-axes then give

$$F_y = m\,(\dot{v} + ru - pw - g_y) \tag{3.38}$$

$$F_z = m\,(\dot{w} - qu + pv - g_z) \tag{3.39}$$

where g_y and g_z are the gravitational components and F_y, F_z the aerodynamic force components. The rigid body rotational equations are

$$M_x = I_{xx}\dot{p} \tag{3.40}$$

$$M_y = I_{yy}\dot{q} - (I_{zz} - I_{xx})rp \tag{3.41}$$

$$M_z = I_{zz}\dot{r} - (I_{xx} - I_{yy})pq \tag{3.42}$$

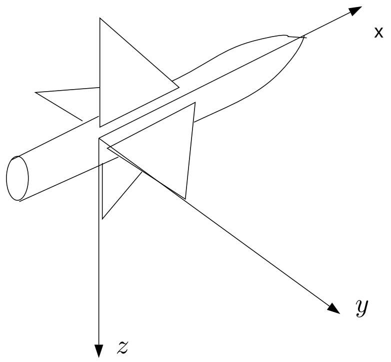


Figure 3.5: Coordinate system of missile

where M_x, M_y and M_z are the aerodynamic torques and I_{xx}, I_{yy} and I_{zz} the moments of inertia. Introducing the state variables

$$x_1 = w/u, \quad x_2 = v/u, \quad x_3 = p, \quad x_4 = q, \quad x_5 = r \tag{3.43}$$

neglecting gravitation and assuming $I_{yy} = I_{zz} = I$ gives

$$\dot{x}_1 = \qquad x_4 - x_3 x_2 + \tilde{C}_z \tag{3.44}$$

$$\dot{x}_2 = \qquad -x_5 + x_3 x_1 + \tilde{C}_y \tag{3.45}$$

$$\dot{x}_3 = \qquad \tilde{C}_\ell \tag{3.46}$$

$$\dot{x}_4 = \qquad I_q x_3 x_5 + \tilde{C}_m \tag{3.47}$$

$$\dot{x}_5 = \qquad -I_q x_3 x_4 + \tilde{C}_n \tag{3.48}$$

where

$$\tilde{C}_Y = \frac{F_y}{mu}, \quad \tilde{C}_Z = \frac{F_z}{mu}$$

$$\tilde{C}_\ell = \frac{M_x}{I_{xx}}, \quad \tilde{C}_m = \frac{M_y}{I_{yy}}, \quad \tilde{C}_n = \frac{M_z}{I_{zz}}, \quad I_q = \frac{I - I_{xx}}{I}$$

Introduce

$$\tilde{C} = \left(\tilde{C}_Z, \tilde{C}_Y, \tilde{C}_\ell, \tilde{C}_m . \tilde{C}_n\right)^T$$

The elements of $\tilde{C}$ depend on the state variables and control signals. We assume that the missile is controlled by the three control signals

$$u_1 = \delta_e, \quad u_2 = \delta_r, \quad u_3 = \delta_a$$

corresponding to elevator, rudder and ailerons respectively. We assume for simplicity that the control acts linearly to give

$$\tilde{C} = F(x) + g(x)u$$

Here F and g depend mainly on x_1 and x_2, while the dependence on the angular velocities x_3, x_4, x_5 is much weaker. A typical structure for g is

$$g(x) = \begin{bmatrix} g_{11}(x) & 0 & 0 \\ 0 & g_{22}(x) & 0 \\ 0 & 0 & g_{33}(x) \\ g_{41}(x) & 0 & 0 \\ 0 & g_{52}(x) & 0 \end{bmatrix} \tag{3.49}$$

with g_{11} much smaller than g_{41} and g_{22} much smaller than g_{52}. The model of the missile has the form

$$\dot{x} = f(x) + g(x)u$$

with f given by

$$f(x) = \begin{bmatrix} x_4 - x_3x_2 + F_1(x) \\ -x_5 + x_3x_1 + F_2(x) \\ F_3(x) \\ I_q x_3 x_5 + F_4(x) \\ -I_q x_3 x_4 + F_5(x) \end{bmatrix} \tag{3.50}$$

The simplest situation is when $g(x)$ is actually constant. Then the transformation

$$z_1 = x_1 - \gamma_1 x_4 \tag{3.51}$$
$$z_2 = x_2 - \gamma_2 x_5 \tag{3.52}$$
$$z_3 = x_3 \tag{3.53}$$
$$z_4 = f_1(x) - \gamma_1 f_4(x) \tag{3.54}$$
$$z_5 = f_2(x) - \gamma_2 f_5(x) \tag{3.55}$$

with $\gamma_1 = g_{11}/g_{41}$ and $\gamma_2 = g_{22}/g_{52}$ gives

$$\dot{z}_1 = z_4 \tag{3.56}$$
$$\dot{z}_2 = z_5 \tag{3.57}$$
$$\dot{z}_3 = f_3(x) + g_{33}u_3 \tag{3.58}$$
$$\dot{z}_4 = (f_1(x) - \gamma_1 f_4(x))_x(f(x) + gu) \tag{3.59}$$
$$\dot{z}_5 = (f_2(x) - \gamma_2 f_5(x))_x(f(x) + gu) \tag{3.60}$$

With z_1, z_2 and z_3 regarded as outputs, this system has relative degree $(2, 2, 1)$ and feedback linearization is possible. If g depends on x the situation is more complicated. It is then necessary that g_1, g_2 and g_3 depend on x in such a way that they are involutive. ∎

3.7 Exercises.

3.1 Suppose the system of Exercise 3.5 has the output

$$y = \sqrt{1 + x_2}$$

What is the relative degree? What is the zero dynamics?

3.2 Consider the bilinear system

$$\begin{aligned} \dot{x}_1 &= -x_1 + u \\ \dot{x}_2 &= -2x_2 + x_1 u \\ y &= x_1 + x_2 \end{aligned}$$

What is the relative degree? What is the zero dynamics?

3.3 Consider the heat exchanger of example 1.2.

$$\frac{d}{dt}(CT) = qcT_0 - qcT + \kappa(T_h - T)$$

Let the state variables be $x_1 = T$, $x_2 = q$ and $T_h = x_3$. Suppose x_2 and x_3 are controlled from the inputs u_1 and u_2 with some lag due to time constants in the control actuators. If $T_0 = 0$, $c/C = 1$ and $\kappa/C = 1$, then the system is described by

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_3 - x_2 x_1 \\ \dot{x}_2 &= -x_2 + u_1 \\ \dot{x}_3 &= -x_3 + u_2 \\ y_1 &= x_1 \\ y_2 &= x_2 \end{aligned}$$

What is the relative degree? Describe a controller that linearizes and decouples the system.

3.4 Transform the system

$$\begin{aligned} \dot{x}_1 &= x_1^2 + x_2 \\ \dot{x}_2 &= -x_2 + u \end{aligned}$$

into controller form. Compute a state feedback which gives linear dynamics with poles in -1 and -2.

3.5 Transform the system

$$\begin{aligned} \dot{x}_1 &= 1 - \sqrt{1 + x_1} + u \\ \dot{x}_2 &= \sqrt{1 + x_1} - \sqrt{1 + x_2} \end{aligned}$$

into controller form. Compute a state feedback giving linear dynamics with poles in -2 and -3.

3.6 Transform the system

$$\begin{aligned} \dot{x}_1 &= \sin x_2 \\ \dot{x}_2 &= \sin x_3 \\ \dot{x}_3 &= u \end{aligned}$$

to controller form. In what region of the state space is the transformation well defined ?

Chapter 4

Nonlinear observers

The use of exact feedback linearization described in the previous chapter requires state feedback. This is also true of many other techniques that we will consider in the coming chapters. Since all state variables are seldom measured, there is a need to reconstruct the state from some measured output y. A device which does this is called an *observer*. We will assume that the system has the form

$$\dot{x} = f(x, u), \quad y = h(x) \tag{4.1}$$

where x, u and y are vectors with dimensions n, m and p respectively.

4.1 A straightforward observer

Recall that a linear system

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

has a natural observer of the form

$$\dot{\hat{x}} = A\hat{x} + Bu + K\,(y - C\hat{x}) \tag{4.2}$$

The observer error $\tilde{x} = x - \hat{x}$ satisfies

$$\dot{\tilde{x}} = (A - KC)\tilde{x} \tag{4.3}$$

and for an observable pair A, C the observer gain K can be chosen to give $A - KC$ arbitrary eigenvalues.

A natural generalization of this observer to the nonlinear case (4.1) is to use

$$\dot{\hat{x}} = f(\hat{x}, u) + K(\hat{x})(y - h(\hat{x})) \tag{4.4}$$

There are many strategies for choosing the gain K. Let the system (4.1) have an equilibrium at the origin:

$$f(0, 0) = 0, \quad h(0) = 0 \tag{4.5}$$

with linearization

$$A = f_x(0,0), \quad C = h(0) \tag{4.6}$$

where the pair A, C is detectable. Then K can be chosen so that $A - K(0)C$ has eigenvalues in the left half plane, guaranteeing an observer which converges in a neighborhood of the origin.

A popular approach is the *extended Kalman filter*. Here the system (4.1) is linearized around $\hat{x}$ at each instance of time and K is chosen from a Kalman filter design based on that linearization. The extended Kalman filter works well in many applications, but there are few hard results on convergence and performance.

4.2 Observers based on differentiation

Consider to begin with a system without inputs.

$$\dot{x} = f(x), \quad y = h(x) \tag{4.7}$$

If $p = n$, i.e. the number of outputs equals the number of states, then one can construct a static observer by solving the nonlinear system of equations

$$h(x) = y \tag{4.8}$$

with respect to x, at each time instant. The question of solvability of this system of equations then arises. The basic mathematical fact that is used is the following:

Theorem 4.1 The Implicit Function Theorem. Consider the system of equations

$$F(x,y) = 0 \tag{4.9}$$

where $x \in R^n$, $y \in R^m$ and F is continuously differentiable, $R^{n+m} \rightarrow R^n$. Let x_o, y_o be such that

$$F(x_o, y_o) = 0, \quad F_x(x_o, y_o) \text{ nonsingular}$$

Then for all y in a neighborhood of y_o (4.9) has a solution

$$x = \phi(y), \quad x_o = \phi(y_o)$$

for some continuously differentiable function ϕ.

From the implicit function theorem it follows that (4.8) can be solved at least locally if the Jacobian h_x is nonsingular.

Now consider the case of (4.7) under the more realistic assumption that $p < n$. One can then get a system of equations with as many equations as unknowns by considering also derivatives

$$\dot{y}_i = L_f h_i(x), \;\; \ddot{y}_i = L_f^2 h_i(x), \ldots, y_i^{(\sigma_i - 1)} = L_f^{\sigma_i - 1} h_i(x)$$

Suppose we can find integers σ_1,...,σ_p, $\sigma_1 + \sigma_2 + \cdots + \sigma_p = n$ such that the row vectors

$$\begin{array}{llll} h_{1,x}, & (L_f h_1)_x, & \dots & (L_f^{\sigma_1 - 1} h_1)_x \\ \vdots & & & \vdots \\ h_{p,x}, & (L_f h_p)_x, & \dots & (L_f^{\sigma_p - 1} h_p)_x \end{array}$$

are linearly independent for some x_o. Then by the implicit function theorem we can solve the corresponding system of equations

$$\begin{array}{llll} y_1 = h_1, & \dot{y}_1 = L_f h_1, & \dots & y^{(\sigma_1 - 1)} = L_f^{\sigma_1 - 1} h_1 \\ & & & \vdots\vdots \\ y_p = h_p, & \dot{y}_p = L_f h_p, & \dots & y^{(\sigma_p - 1)} = L_f^{\sigma_p - 1} h_p \end{array} \tag{4.10}$$

in a neighborhood of that x_o. Often there are many ways of choosing the numbers σ_1,..,σ_p that satisfy the linear independence requirements. However, one usually wants to differentiate physical signals as few times as possible. Therefore it is natural to make the choice so that the highest σ_i becomes as small as possible. If this leaves some freedom of choice the second highest σ_i should be as small as possible, and so on. If the choice is made in this way, the numbers σ_1.,,σ_p are called the *observability indices* of the system.

We assume that every index satisfies $\sigma_i \geq 1$, that is we assume that every output signal is used (otherwise we could delete one signal and regard the system as one with fewer outputs). By reordering the output variables we can always get the indices in decreasing order:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 1 \tag{4.11}$$

The indices σ_i give the lengths of the rows in the array (4.10). The same information is given by instead listing the lengths of the columns. We define

$$\ell_i = \text{number of outputs that are differentiated } \ell_i \text{ times} \tag{4.12}$$

Since we assume that all outputs are used, we have $\ell_0 = p$. If the σ_i are ordered in decreasing order we have

$$p = \ell_0 \geq \ell_1 \geq \cdots \geq \ell_{\sigma_1 - 1} > 0 \tag{4.13}$$

(Some authors call the ℓ_i the observability indices.)

An observer based on differentiation performs the calculations

1. Compute approximations of the derivatives

$$\begin{array}{llll} y_1, & \dot{y}_1, & \dots & y^{(\sigma_1 - 1)} \\ \vdots & & & \vdots \\ y_p, & \dot{y}_p, & \dots & y^{(\sigma_p - 1)} \end{array}$$

This can for instance be done using differentiating filters of the form

$$\frac{s}{1+sT}$$

where the filter time constant T has to be chosen as a compromise between the desire to get an accurate derivative and the need to filter out high frequency disturbances. Often the output is measured at discrete points of time and then the problem of approximating the derivatives can be approached through numerical differentiation, based on difference schemes.

2. Solve the equations (4.10). There is no general way of performing this step. For many physical systems the structure is such that an explicit solution is possible. Otherwise a numerical solution has to be sought. The non-singularity of the Jacobian guarantees that a Newton iteration converges if the initial guess is good enough. If the equations consist of polynomials, then Gröbner basis techniques might be used.

Differentiation and input signals

So far we have looked at systems without inputs. If we include an input

$$\dot{x} = f(x) + b(x)u, \quad y = h(x) \tag{4.14}$$

the situation becomes more complex, since the possibility of calculating x from u and y will in general depend on the choice of input. Take for example the system

$$\begin{aligned}\dot{x}_1 &= x_2 u\\ \dot{x}_2 &= 0\\ y &= x_1\end{aligned}$$

It is clear that if $u = 0$ there is no way to determine x_2, while for a nonzero u, x_1 and x_2 can be calculated from

$$x_1 = y, \quad x_2 = \dot{y}/u$$

To study this problem in more detail we look at single-input-single-output systems of the form (4.14). To make the study easier we introduce the variables

$$z_1 = h(x), \quad z_2 = (L_f h)(x), \ldots, z_n = (L_f^{n-1} h)(x) \tag{4.15}$$

or in vector notation

$$z = \begin{bmatrix} h(x) \\ (L_f h)(x) \\ \vdots \\ (L_f^{n-1} h)(x) \end{bmatrix} = \Phi(x) \tag{4.16}$$

Proposition 4.1 If the Jacobian $\Phi_x(x_o)$ is nonsigular then the variable change (4.16) is invertible in a neighborhood of x_o.

Proof. Follows directly from the implicit function theorem. ∎

In the new variables the dynamics becomes simple.

$$\begin{aligned}
&\dot{z}_1 = L_f h + L_b h \ u = z_2 + g_1(z)u, \quad \text{where } g_1(z) = L_b h \\
&\dot{z}_2 = L_f^2 h + L_b L_f h \ u = z_3 + g_2(z)u, \quad \text{where } g_2(z) = L_b L_f h \\
&\ \ \vdots \\
&\dot{z}_n = L_f^n h + L_b L_f^{n-1} h \ u = \phi(z) + g_n(z)u, \quad \text{where } \phi(x) = L_f^n h, \ g_n(z) = L_b L_f^{n-1} h
\end{aligned}$$

The system description in the z-variables is thus

$$\dot{z} = \begin{bmatrix} z_2 \\ \vdots \\ z_n \\ \phi(z) \end{bmatrix} + \begin{bmatrix} g_1(z) \\ \vdots \\ g_{n-1}(z) \\ g_n(z) \end{bmatrix} u \tag{4.17}$$

This form has a number of interesting properties.

Proposition 4.2 If $u = 0$ then z in (4.17) can be computed from y, $\dot{y}$,..,$y^{(n-1)}$.

Proof. Equation (4.17) directly gives

$$z_1 = y, \quad z_2 = \dot{y}, \ldots z_n = y^{(n-1)}$$

∎

If (4.17) has a special structure then it is possible to compute z for any choice of u (provided u is known).

Proposition 4.3 Let (4.17) have the form

$$\dot{z} = \begin{bmatrix} z_2 \\ z_3 \\ \vdots \\ z_n \\ \phi(z) \end{bmatrix} + \begin{bmatrix} g_1(z_1) \\ g_2(z_1, z_2) \\ \vdots \\ g_{n-1}(z_1, \ldots, z_{n-1}) \\ g_n(z) \end{bmatrix} u \tag{4.18}$$

Then z can be computed from u, y and their derivatives no matter how u is chosen.

Proof. From (4.18) it follows that, for each j

$$z_j = \dot{z}_{j-1} - g_{j-1}(z_1, \ldots, z_{j-1})u$$

Since each z_j only depends on the variables with lower indices, the z_j can be calculated recursively, starting with $z_1 = y$. ∎

4.3 High gain observers

In the previous section z was computed using derivatives of y and u (or their approximations). With the structure (4.18) there is also a possibility of using the observer structure (4.4) to avoid explicit calculations of derivatives. If we introduce the matrices

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & 0 \\ \vdots & & & & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \tag{4.19}$$

the system can be written

$$\dot{z} = Az + B\phi(z) + g(z)u, \quad y = Cz \tag{4.20}$$

The natural observer is then

$$\dot{\hat{z}} = A\hat{z} + B\phi(\hat{z}) + g(\hat{z})u + K(y - C\hat{z}) \tag{4.21}$$

The observer error $\tilde{z} = z - \hat{z}$ then satisfies

$$\dot{\tilde{z}} = (A - KC)\tilde{z} + B(\phi(z) - \phi(\hat{z})) + (g(z) - g(\hat{z}))\, u \tag{4.22}$$

where the matrix

$$A - KC$$

has the structure

$$A - KC = \begin{bmatrix} -k_1 & 1 & 0 & \dots & 0 \\ -k_2 & 0 & 1 & & 0 \\ \vdots & & & \ddots & 0 \\ \vdots & & & & 1 \\ -k_n & 0 & 0 & \dots & 0 \end{bmatrix} \tag{4.23}$$

To simplify things we consider the situation when $u = 0$. Using the notation $\delta = (\phi(z) - \phi(\hat{z}))$ we can write

$$\dot{\tilde{z}} = (A - KC)\tilde{z} + B\delta$$

We see that the observer can be regarded as a linear system with the contribution from the nonlinearity as an external signal. The transfer function from δ to $\tilde{z}$ is given by

$$(sI - A + KC)^{-1}B = \frac{1}{s^n + k_1 s^{n-1} + \cdots + k_n} \begin{bmatrix} 1 \\ k_1 + s \\ \vdots \\ s^{n-1} + k_1 s^{n-2} + \cdots + k_{n-1} \end{bmatrix} \tag{4.24}$$

Now suppose K is chosen in the following way

$$k_1 = \beta_1/\epsilon, k_2 = \beta_2/\epsilon^2, \dots, k_n = \beta_n/\epsilon^n$$

where β_1,..,β_n are chosen such that

$$s^n + \beta_1 s^{n-1} + \cdots + \beta_n = 0$$

has roots strictly in the left half plane. Using these relations in (4.24) gives the following tranfer function from δ to $\tilde{z}$.

$$(sI - A + KC)^{-1}B =$$

$$\frac{1}{(\epsilon s)^n + \beta_1(\epsilon s)^{n-1} + \cdots + \beta_n} \begin{bmatrix} \epsilon^n \\ \epsilon^{n-1}(\beta_1 + (\epsilon s)) \\ \vdots \\ \epsilon((\epsilon s)^{n-1} + \beta_1(\epsilon s)^{n-2} + \cdots + \beta_{n-1}) \end{bmatrix} \tag{4.25}$$

Since all the transfer functions go to zero as ϵ tends to zero the influence of the nonlinear terms on the estimation error becomes negligible. Intuitively this observer should therefore work well for small ϵ irrespective of the form of the nonlinearity. To get precise results we need tools from stability theory that will be presented in the following chapters. We will then return to the question of stability of the high gain observer.

There are some obvious disadvantages of the high gain observer:

- The measurement signal y is multiplied by the coefficients $k_i = \beta_i/\epsilon_i$ that tend to infinity as ϵ tends to zero. Any measurement error will therefore be amplified by a factor that can become very large.
- Since s in (4.25) always occurs as ϵs the parameter ϵ will act as a frequency scaling so that the bandwidths of the transfer functions go to infinity as ϵ goes to zero. The observer could therefore be sensitive to modeling errors or noise at high frequencies.

4.4 Observers based on state transformations

There is one form of the system equations that is nice for proving convergence of the observer, namely

$$\dot{x} = Ax + f(u, y), \quad y = Cx \tag{4.26}$$

With the observer

$$\dot{\hat{x}} = A\hat{x} + f(u, y) + K(y - C\hat{x}) \tag{4.27}$$

the estimation error $\tilde{x} = x - \hat{x}$ satisfies

$$\dot{\tilde{x}} = (A - KC)\tilde{x} \tag{4.28}$$

It is thus possible to get global convergence of the observer error, provided the pair A, C is detectable. The form (4.26) is very special, but there is the possibility of transforming a wider class of systems into that form. To begin with we will look at systems without inputs and try to find a transformation $x = X(\xi)$ which transforms the system

$$\dot{x} = f(x), \quad y = h(x) \tag{4.29}$$

into the form

$$\dot{\xi} = \tilde{f}(\xi), \qquad y = \tilde{h}(\xi) \tag{4.30}$$

$$\tilde{f}(\xi) = A\xi + b(y), \quad \tilde{h}(\xi) = C\xi \tag{4.31}$$

Let Y denote the inverse transformation: $\xi = Y(x) = X^{-1}(x)$. We assume that the system (4.29) is characterized by the observability indices σ_i and the parameters ℓ_i satisfying (4.11), (4.12), (4.13). We use the notation $\sigma = \sigma_1$. Introduce the observability matrix

$$Q = \begin{bmatrix} h_{1,x} \\ \vdots \\ h_{\ell_0,x} \\ (L_f h_1)_x \\ \vdots \\ (L_f h_{\ell_1}) \\ \vdots \\ (L_f^\sigma h_1)_x \\ \vdots \\ (L_f^\sigma h_{\ell_{\sigma-1}})_x \end{bmatrix} \tag{4.32}$$

It follows from the definition of the observability indices that Q is nonsingular in a neighborhood of some point x_o.

To simplify the calculations it is desirable to pick A and C matrices which are simple. In the single output case every observable pair A, C can be transformed into observer canonical form

$$A = \begin{bmatrix} -a_1 & 1 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & & 0 \\ \vdots & & & & \vdots \\ -a_{n-1} & 0 & \dots & 0 & 1 \\ -a_n & 0 & \dots & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

If the linear terms depending on $\xi_1 = y$ are included in $b(y)$ the system (4.30) then takes the form

$$\begin{aligned} \dot{\xi}_1 &= b_1(\xi_1) + \xi_2 \\ \dot{\xi}_2 &= b_2(\xi_1) + \xi_3 \\ &\vdots \\ \dot{\xi}_n &= b_n(\xi_1) \\ y &= \xi_1 \end{aligned} \tag{4.33}$$

which can be regarded as a nonlinear observer form.

In the multi-output case an observable pair A, C can be transformed into

$$A = \begin{bmatrix} \times & E_1 & 0 & \cdots & \\ \times & 0 & E_2 & & \\ & & & \ddots & \\ & & & 0 & E_{\sigma-1} \\ \times & \cdots & & \cdots & 0 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & 0 & \cdots & 0 \end{bmatrix} \tag{4.34}$$

where the block in position i, j has dimensions $\ell_{i-1} \times \ell_{j-1}$. E_i consists of the first ℓ_{i-1} rows and the first ℓ_i columns of an identity matrix, C_1 is an inverible matrix. The blocks marked × depend only on the first ℓ_0 states, i.e. on y. If they are included in $b(y)$, then the linear part is described by

$$A = \begin{bmatrix} 0 & E_1 & 0 & \cdots & \\ 0 & 0 & E_2 & & \\ & & & \ddots & \\ & & & 0 & E_{\sigma-1} \\ 0 & \cdots & & \cdots & 0 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & 0 & \cdots & 0 \end{bmatrix} \tag{4.35}$$

The pair A, C then constitutes a so called *condensed Brunovsky canonical form*. For example, if $\sigma_1 = 3$ and $\sigma_2 = 1$, then $\ell_0 = 2$, $\ell_1 = 1$, $\ell_2 = 1$ and

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} \times & \times & 0 & 0 \\ \times & \times & 0 & 0 \end{bmatrix}$$

If $C_1 = I$ the canonical form is called (dual) Brunovsky form. We will see below that this case gives (fairly) simple calculations. Unfortunately some systems can only be transformed into (4.30) by choosing a C_1 different from the identity matrix.

Now consider the properties of the transformation $x = X(\xi)$. Define the Jacobian

$$J(\xi) = X_\xi(\xi) \tag{4.36}$$

By differentiating the relation $x = X(\xi)$ we get

$$\dot{x} = J(\xi)\dot{\xi} = J(\xi)\tilde{f}(\xi) \tag{4.37}$$

Since we also have $\dot{x} = f(x)$, we have the following relation between the right hand sides of the equations

$$f(X(\xi)) = J(\xi)\tilde{f}(\xi) \tag{4.38}$$

Now suppose we have a different pair of vector fields that are transformed in the same way:

$$g(X(\xi)) = J(\xi)\tilde{g}(\xi) \tag{4.39}$$

We have the following relation between their Lie brackets.

Proposition 4.4 Let the vector fields f, g be transformed as in (4.38), (4.39). Then their Lie Brackets are related as

$$[f,g] = J[\tilde{f},\tilde{g}] \tag{4.40}$$

Here the Lie bracket in the left hand side is evaluated using differentiation with respect to x, while the bracket in the right hand side uses differentiation with respect to ξ.

Proof. Exercise. ■

If g is a matrix with columns g_i, $g = (g_1,\dots,g_k)$ we can interpret the Lie bracket columnwise:

$$[f,g] = [f,(g_1,\dots,g_k)] = ([f,g_1],\dots,[f,g_k]) \tag{4.41}$$

Proposition 4.4 then remains true. We can now use this formalism to describe the transformation from (4.29) to (4.30). Take $\tilde{g}$ to be the unit matrix in the ξ coordinates. We then have

$$[f,g] = J[\tilde{f},\tilde{g}] = J[\tilde{f},I] = -J\tilde{f}_\xi = -J(A+b_\xi) \tag{4.42}$$

Since $g(X(\xi)) = J(\xi)\tilde{g}(\xi) = J(\xi)$ we have, rewriting in the x-coordinates

$$[f(x),J(Y(x))] = -J(A+b_\xi(Y(x))) \tag{4.43}$$

Now partition J into σ blocks $J = (J_1,J_2,\dots,J_\sigma)$ where the number of columns in each block is $\ell_0,\ell_1,..\ell_{\sigma-1}$ respectively. Interpreting (4.43) for each block column gives

$$-[f,J_1] = Jb_{\xi_1} \tag{4.44}$$

$$-[f,J_k] = J_{k-1}E_{k-1}, \quad k = 2,\dots,\sigma \tag{4.45}$$

where b_{ξ_1} is the Jacobian of b with respect to the first $\ell_0 = p$ components of ξ (since b depends only on the outputs, it depends only on those variables). We get further relations by considering the output equations

$$y = h(x), \quad y = \tilde{h}(\xi) = C\xi$$

Differentiating the output we have

$$\begin{aligned} h(x) &= y = \tilde{h}(\xi) \\ (L_f h)(x) &= \dot{y} = (L_{\tilde{f}}\tilde{h})(\xi) \\ &\vdots \\ (L_f^{\sigma-1}h)(x) &= y^{(\sigma-1)} = (L_{\tilde{f}}^{\sigma-1}\tilde{h})(\xi) \end{aligned}$$

Differentiating with respect to ξ gives

$$\begin{bmatrix} h_x \\ (L_f h)_x \\ \vdots \\ (L_f^{\sigma-1}h)_x \end{bmatrix} J = \begin{bmatrix} \tilde{h}_\xi \\ (L_{\tilde{f}}\tilde{h})_\xi \\ \vdots \\ (L_{\tilde{f}}^{\sigma-1}\tilde{h})_\xi \end{bmatrix}$$

Evaluating the right hand side finally gives

$$\begin{bmatrix} h_x \\ (L_f h)_x \\ \vdots \\ (L_f^{\sigma-1} h)_x \end{bmatrix} J = \begin{bmatrix} C_1 & 0 & 0 & \dots & 0 \\ * & C_1 E_1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \\ * & * & \dots & & C_1 E_1 \cdots E_{\sigma-1} \end{bmatrix} \tag{4.46}$$

where the elements marked "*" depend on b. From (4.46) and (4.45) it is possible to calculate the desired coordinate change if it exists. First we take a look at the single variable case.

Observer form for single output systems

For $p = 1$ we can pick out the last column of equation (4.46) to get

$$\begin{bmatrix} h_x \\ (L_f h)_x \\ \vdots \\ (L_f^{\sigma-1} h)_x \end{bmatrix} J_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{4.47}$$

For the single output case the matrix to the left is just Q, which we assumed to be nonsingular. This equation can therefore be solved to give J_n. The other columns of J can then be calculated successively from (4.45) which for the single output case becomes

$$J_{k-1} = -[f, J_k], \quad k = n, \dots, 2 \tag{4.48}$$

After solving for J we can calculate the transformation from

$$Y_\xi(x) = J^{-1}(x) \tag{4.49}$$

provided this system is integrable, that is if $J^{-1}(x)$ actually is the Jacobian of some vector valued function. We have the following theorem.

Theorem 4.2 The problem of transforming (4.29) into (4.30) can be solved for the scalar output case if and only if

1. The matrix Q is nonsingular.
2. $Y_x(x) = J^{-1}(x)$ with the columns of J given by (4.47), (4.48) is integrable.

The solution is obtained by computing $\xi = Y(x)$ from $Y_\xi(x) = J^{-1}(x)$.

Proof. Our calculations above have shown that the conditions 1. and 2. are necessary. The sufficiency follows as a special case of the multivariable case discussed below. ∎

To illustrate the importance of the integrability condition we look at two examples.

Example 4.1 Consider the system

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= x_1 x_2\\ y &= x_1\end{aligned}$$

Since $Q = I$ we immediately get $J_2 = (0\ \ 1)^T$. J_1 is then calculated from

$$J_1 = -[f, J_2] = f_x J_2 = \begin{bmatrix} 0 & 1 \\ x_2 & x_1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ x_1 \end{bmatrix}$$

giving

$$J = \begin{bmatrix} 1 & 0 \\ x_1 & 1 \end{bmatrix}, \quad J^{-1} = \begin{bmatrix} 1 & 0 \\ -x_1 & 1 \end{bmatrix}$$

The first part of the coordinate change is already given by $y = x_1 = \xi_1$. The second part has to satisfy

$$\begin{aligned}\frac{\partial \xi_2}{\partial x_1} &= -x_1\\ \frac{\partial \xi_2}{\partial x_2} &= 1\end{aligned}$$

The first equation gives

$$\xi_1 = -\frac{x_1^2}{2} + \phi(x_2)$$

where ϕ is an arbitrary function. Differentiating and using the second equation gives

$$\phi'(x_2) = 1 \Rightarrow \phi_2(x_2) = x_2 + k$$

for some constant k. Picking $k = 0$ gives the coordinate transformation

$$\begin{aligned}\xi_1 &= x_1\\ \xi_2 &= -\frac{x_1^2}{2} + x_2\end{aligned}$$

The system description in the new coordinates becomes

$$\begin{aligned}\dot{\xi}_1 &= \frac{\xi_1^2}{2} + \xi_2\\ \dot{\xi}_2 &= 0\\ y &= \xi_1\end{aligned}$$

■

Example 4.2 Now consider the system

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= x_1^2 + x_2^2\\ y &= x_1\end{aligned}$$

Performing calculations similar to those of the previous example we get

$$J = \begin{bmatrix} 1 & 0 \\ 2x_2 & 1 \end{bmatrix}, \quad J^{-1} = \begin{bmatrix} 1 & 0 \\ -2x_2 & 1 \end{bmatrix}$$

The second ξ-coordinate then has to satisfy

$$\frac{\partial \xi_2}{\partial x_1} = -2x_2$$
$$\frac{\partial \xi_2}{\partial x_2} = 1$$

Solving the first equation gives

$$\xi_2 = -2x_1x_2 + \phi(x_2)$$

for some arbitrary function ϕ_2. Differentiating with respect to x_2 and using the second equation gives

$$\frac{\partial \xi_2}{\partial x_2} = -2x_1 + \phi'(x_2) = 1$$

This relation can not be satisfied for any choice of ϕ depending only on x_2 and the problem is thus not solvable. ■

The integrability can be checked without actually computing the transformation $x = X(\xi)$.

Proposition 4.5 A necessary and sufficient condition for integrability of $Y_x(x) = J^{-1}(x)$ is that

$$[(ad^j f, J_n), (ad^k f, J_n)] = 0, \quad j = 1, \ldots, n-1, \quad k = 1, \ldots, n-1 \qquad (4.50)$$

Proof. We give only the necessary part of the proof. We have $Y_x(x)J(x) = I$. It follows that $Y_xJ_k = e_k$, where e_k the k:th column of the identity matrix. From Proposition 4.4 we get $Y_x[J_i, J_k] = [e_i, e_k] = 0$ implying $[J_i, J_k] = 0$. Since $J_{n-1} = -(ad^1 f, J_n)$, $J_{n-2} = (ad^2 f, J_n)$ etc., the result follows. ■

Remark Since Lie brackets satisfy the Jacobi identity:

$$[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0 \qquad (4.51)$$

it is not necessary to check all the brackets of (4.50). Some of them can be deduced from the others.

Observer form for multi-output systems with $C_1 = I$

For multivariable systems we first consider the case where the transformation can actually be made to dual Brunovsky canonical form so that $C_1 = I$. Equation (4.46) then takes the form

$$\begin{bmatrix} h_x \\ (L_f h)_x \\ \vdots \\ (L_f^{\sigma-1} h)_x \end{bmatrix} J = \begin{bmatrix} I & 0 & 0 & \ldots & 0 \\ * & \bar{E}_1 & 0 & \ldots & 0 \\ \vdots & & \ddots & & \\ * & * & \ldots & & \bar{E}_{\sigma-1} \end{bmatrix} \qquad (4.52)$$

where $\bar{E}_i$ is a $p \times \ell_i$ matrix whose first ℓ_i rows form a unit matrix. If we look at the rows corresponding to $L_f^j h$ we have the equation

$$\begin{bmatrix} (L_f^j h_t j)_x \\ (L_f^j h_{bj})_x \end{bmatrix} J_j = \begin{bmatrix} I \\ 0 \end{bmatrix} \tag{4.53}$$

where h_{tj} contains the first ℓ_j rows of h and h_{bj} the remaining ones. From the definition of the observability indices it follows that $(L_f^j h_{bj})_x$ must be linearly dependent on the elements above it in the matrix forming the left hand side of (4.52) (otherwise it would be possible to lower one of the σ_i). However, from (4.53) it follows that $(L_f^j h_{bj})_x$ can not be linearly dependent on $(L_f^j h_{tj})_x$. We then have

$$(L_f^J h_{bj})_x \in \operatorname{span}\{h_x, \ldots, (L_f^{j-1} h)_x\} \tag{4.54}$$

One can also show the reverse: when this criterion is satisfied and it is possible to achieve the condensed Brunovsky form, then it is actually possible to achieve $C_1 = I$, that is the dual Brunovsky form. Assuming (4.54) to be true we can proceed with the solution of (4.52). Introduce

$$Q_j = \begin{bmatrix} h_x \\ \vdots \\ (L_f^{j-1} h)_x \end{bmatrix}, \quad \bar{\bar{E}}_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \bar{E}_{j-1} \end{bmatrix} \tag{4.55}$$

Take a block structure $J = (J_1, \ldots, J_\sigma)$ which is compatible with the blocks of the right hand side of (4.52). If we pick out those equations in (4.52) that have known right hand sides, (i.e. independent of b) and combine them with (4.45) we get the system

$$Q_\sigma J_\sigma = \bar{\bar{E}}_\sigma \tag{4.56}$$

$$Q_j J_j = \bar{\bar{E}}_j, \quad J_j E_j = -[f, J_{j+1}], \quad j = \sigma - 1, \ldots, 1 \tag{4.57}$$

This system can be solved using generalized inverses.

Lemma 4.1 (Rao and Mitra, 1971) The individually consistent matrix equations $DY = E$ and $YF = H$ have a common solution if and only if $EF = DH$. The general solution is then

$$Y = D^- E + HF^- - D^- DHF^- + (I - D^- D)Z(I - FF^-) \tag{4.58}$$

where D^-, F^- are arbitrary generalized inverses (i.e. satisfying $DD^-D = D$, $FF^-F = F$) and Z is an arbitrary matrix of appropriate dimension.

Using this lemma it is easy to show the following result.

Lemma 4.2 Assume rank $Q = n$. The set of equations (4.56), (4.57) are solvable if (4.54) is satisfied and the solution is given by

$$J_\sigma = Q_\sigma^- \bar{\bar{E}}_\sigma, \quad J_j = -[f, J_{j+1}]E_j^T + Q_j^- \bar{\bar{E}}_j(I - E_j E_j^T)$$
$$+ (I - Q_j^- Q_j) Z_j (I - E_i E_i^T) \tag{4.59}$$

Proof. Some Lie derivative calculations show that the condition $EF = DH$ of Lemma 4.1 is satisfied by (4.57). The formula (4.59) for the solution then follows from that lemma on noting that $E_j^- = (E_j^T E_j)^{-1} E_j^T = E_j^T$. ■

Observer form for multi-output systems with general C_1

For a system in the form (4.35) it is possible to get the dual Brunovsky form by introducing a transformed output $\tilde{y} = C_1^{-1} y$. Of course C_1 is not known before the transformation $x = X(\xi)$ has been computed. It is however possible to deduce C_1 from the condition that the transformed output equation

$$\tilde{y} = C_1^{-1} y = C_1^{-1} h(x)$$

should satisfy (4.54). If for instance

$$(L_f^j h_{bj})_x = M(L_f^j h_{tj})_x + \text{span}\{h_x, \dots, (L_f^{j-1} h)x\}$$

then

$$C_1^{-1} = \begin{bmatrix} I & 0 \\ -M & I \end{bmatrix} \tag{4.60}$$

removes the dependence on $(L_f^j h_{tj})_x$. Working through $j = 1, \dots, \sigma$ one can try to satisfy (4.54) by forming C_1 as a product of matrices of the form (4.60). It turns out that this is possible precisely when

$$(L_f^j h_{bj})_x \in \text{span}_R (L_f^j h_{tj})_x + \text{span}\{h_x, \dots, (L_f^{j-1} h)x\}, \quad j = 0, \dots, \sigma - 1 \tag{4.61}$$

where "span" denotes linear combinations with coefficients that are functions of x and "span$_R$" denotes linear combinations with constant coefficients. We have then finally arrived at the general result.

Theorem 4.3 The problem of transforming (4.29) into (4.30) is solvable if and only if
(a) There exist observability indices satisfying (4.11) with $\sigma_1 + \dots + \sigma_p = n$ such that Q has rank n.
(b) (4.61) is satisfied.
(c) The matrix J computed from Lemma 4.2 is integrable.

The solution is then given by $X_\xi = J$.

Proof. (sketch of sufficiency) Using transformations of the form (4.60) the problem is reduced to the case $C_1 = I$. Lemma 4.2 then shows that (4.56), (4.57) can be solved. Some further calculations then show that the starred elements of (4.52) come out right (so that b can be chosen). If computed J is integrable, the required transformation $x = X(\xi)$ can then be computed. ■

4.5 Observers based on state and output transformations

Above we discussed the relation between the two Brunovsky forms by introducing a constant output transformation $\tilde{y} = C_1^{-1} y$. Of course this idea can be extended to include general transformations of the outputs of the form $y = \psi(\tilde{y})$. This gives us additional degrees of freedom and should make it possible to satisfy the very restrictive integrability conditions of Theorem 4.3 more often. Computing the conditions that ψ should satsfy turns out to be very cumbersome. To simplify matters somewhat we make an initial state space transformation of the system. Assuming the matrix Q to be nonsingular, we can introduce the new variables

$$\begin{aligned} &z_{11} = y_1, \quad z_{12} = \dot{y}_1, \ldots, z_{1,\sigma_1} = y_1^{(\sigma_1-1)} \\ &\vdots \\ &z_{p1} = y_p, \quad z_{p2} = \dot{y}_p, \ldots, z_{p,\sigma_p} = y_p^{(\sigma_p-1)} \end{aligned} \tag{4.62}$$

and get a locally invertible coordinate change. In these coordinates the state variable description becomes

$$\begin{aligned} y_1 = z_{11}, &\quad \ldots, y_p = z_{p1} \\ \dot{z}_{11} = z_{12}, &\quad \ldots, \dot{z}_{p1} = z_{p2} \\ \vdots &\qquad\quad \vdots \\ \dot{z}_{1\sigma_1} = f_1(z) &\quad \dot{z}_{p\sigma_p} = f_p(z) \end{aligned} \tag{4.63}$$

where $f_i = L_f^{\sigma_i} h_i$. This form is sometimes referred to as *observable form* or *observability form*. Let Ψ be the Jacobian of ψ and define

$$\Psi_{ij} = \frac{\partial y_i}{\partial \tilde{y}_j} \tag{4.64}$$

The fundamental result is then

Theorem 4.4 Consider a system in observable form (4.63). If it can be transformed into (4.30) with A, C in dual Brunovsky form then Ψ has to satisfy

$$\Psi_{ij} = 0, \quad \sigma_i > \sigma_j \tag{4.65}$$

$$\frac{\partial \Psi_{ij}}{\partial y_\ell} = \frac{1}{\sigma_i} \sum_{k=1}^{p} \frac{\partial^2 f_i}{\partial z_{\ell\sigma_i} \partial z_{k2}} \Psi_{kj} \tag{4.66}$$

In particular the quantities

$$\frac{\partial^2 f_i}{\partial z_{\ell\sigma_i} \partial z_{k2}}$$

have to depend only on the output.

Proof. We consider only the very special case $p = 1$, $n = 2$. The general case involves similar but messier calculations. Comparing

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= f(z_1, z_2) \\ y &= z_1 \end{aligned}$$

with

$$\begin{aligned} \dot{\xi}_1 &= b_1(\xi_1) + \xi_2 \\ \dot{\xi}_2 &= b_2(\xi_1) \\ \tilde{y} &= \xi_1 \end{aligned}$$

shows that

$$z_2 = \dot{y} = \Psi \dot{\tilde{y}} = \Psi(b_1(\xi_1) + \xi_2) \tag{4.67}$$

Differentiating this expression with respect to time gives

$$f(z_1, z_2) = \frac{d\Psi}{dy} \dot{y}(b_1(\xi_1) + \xi_2) + \Psi\big(b_1'(\xi_1)(b_1(\xi_1) + \xi_2) + b_2(\xi_1)\big)$$

Substituting from (4.67) gives

$$f(z_1, z_2) = \frac{d\Psi}{dy} \Psi^{-1} z_2^2 + \Psi\big(b_1'(\xi_1)\Psi^{-1} z_2 + b_2(\xi_1)\big)$$

The key observation is now that, since there is a direct relationship between y and $\tilde{y}$, ξ_1 depends only on z_1 and not on z_2. Differentiating twice with respect to z_2 then gives

$$\frac{\partial^2 f(z_1, z_2)}{\partial z_2^2} = 2\frac{d\Psi}{dy}\Psi^{-1}$$

which is (4.66) for the case $n = 2$, $p = 1$. ■

We illustrate the procedure with the rocket example.

Example 4.3 Consider a vehicle that is accelerated with a unit force and has an aerodynamic drag proportional to the square of the velocity. If x is the position and v is the velocity the system description is

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= 1 - v^2 \\ y &= x \end{aligned}$$

where it is assumed that the position is measured. Equation (4.66) becomes

$$\frac{d\Psi}{d\eta} = -\Psi$$

with the solution $e^{-\eta}$. Thus

$$\frac{d\eta}{d\tilde{y}} = e^{-\eta}, \quad \text{or} \quad \frac{d\tilde{y}}{d\eta} = e^{\eta}$$

showing that

$$\tilde{y} = e^{\eta}$$

We now have the task of transforming the system

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= 1 - v^2 \\ \tilde{y} &= e^x \end{aligned}$$

according to the methods of Theorem 4.4. Since

$$\begin{aligned} h_x &= e^x(1 \quad 0) \\ (L_f h)_x &= e^x(v \quad 1) \end{aligned}$$

we have

$$J_2 = e^{-x} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Then

$$[f, J_2] = e^{-x} \begin{pmatrix} -1 \\ v \end{pmatrix}$$

and $[g, [f, g]] = 0$ showing that (4.50) is satisfied. From (4.66) the coordinate change is

$$J^{-1} = e^{x} \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix}$$

We already knew from the relation between the outputs that

$$z_1 = e^x$$

which is the first row of the above realtion. The second row gives

$$\begin{aligned} \frac{\partial \xi_2}{\partial x} &= ve^x \\ \frac{\partial \xi_2}{\partial v} &= e^x \end{aligned}$$

with the solution $\xi_2 = ve^x$. The coordinate change is thus given by

$$\begin{array}{cc} \xi_1 = e^x & x = \ln \xi_1 \\ \xi_2 = ve^x & v = \xi_2/\xi_1 \end{array}$$

In the new variables the system description is then

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \xi_1 \\ \tilde{y} &= \xi_1 \end{aligned}$$

■

4.6 Observers for systems with input signals

The methods for transforming systems without inputs into observer form can be extended to systems with inputs in several ways. Let the system dynamics be described by

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x) \end{aligned} \tag{4.68}$$

Then we can pick a constant control signal u_0 and compute the variable transformations for the system

$$\begin{aligned} \dot{x} &= f(x, u_0) \\ y &= h(x) \end{aligned} \tag{4.69}$$

using the technique of the previous section (if possible). The resulting transformation can then be applied to the system (4.68) to give

$$\begin{aligned} \dot{\xi} &= b(C\xi) + A\xi + \gamma(\xi, u) \\ \tilde{y} &= C\xi \end{aligned} \tag{4.70}$$

where $\gamma(\xi, u_0) = 0$. If it turns out that γ has the form

$$\gamma(\xi, u) = \gamma(C\xi, u) \tag{4.71}$$

then the observer

$$\dot{\hat{xi}} = b(\tilde{y}) + A\hat{\xi} + \gamma(\tilde{y}, u) + K(\tilde{y} - C\hat{\xi}) \tag{4.72}$$

will have linear error dynamics.

Example 4.4 Suppose that we add the control signal in Example 4.3 to get

$$\begin{array}{rcl} \dot{x} &=& v \\ \dot{v} &=& 1 - v^2 + \tilde{u} \\ y &=& x \end{array}$$

where $\tilde{u} = u - 1 - g$. If we apply the varaible transformation

$$\begin{array}{ll} xi_1 = e^x \qquad & x = \ln \xi_1 \\ \xi_2 = ve^x \qquad & v = \xi_2 / z\xi_1 \end{array}$$

which was computed for $u = 1 + g$, the result is

$$\begin{array}{rcl} \dot{\xi}_1 &=& \xi_2 \\ \dot{\xi}_2 &=& \xi_1 + \xi_1 \tilde{u} \end{array}$$

which is of the form (4.70), (4.71). ■

4.7 Closing the loop

Often observers are used in connection with control. Typically one has a system to be controlled

$$\dot{x} = f(x, u), \quad y = h(x) \tag{4.73}$$

and designs a controller $u = k(x)$ so that the closed loop system

$$\dot{x} = f(x, k(x)) \tag{4.74}$$

has desirable properties. If the state x can not be measured an observer

$$\dot{\hat{x}} = f(\hat{x}, u) + K(y - h(\hat{x})) \tag{4.75}$$

is constructed and the control $u = k(\hat{x})$ is used. The closed loop system then becomes

$$\begin{aligned} \dot{x} &= f(x, k(\hat{x})) \\ \dot{\hat{x}} &= f(\hat{x}, k(\hat{x})) + K(h(x) - h(\hat{x})) \end{aligned} \tag{4.76}$$

It would be nice if a good design of the state feedback system (4.74) and a good design of the observer (4.75) would automatically give a good closed loop design (4.76). For linear systems there are such results, often called *separation theorems*, but for nonlinear systems it has turned out to be very difficult to get such results that are practically useful. This means that it is necessary to analyze (4.76) directly. One possibility is to use Lyapunov theory which will be dealt with in the next chapters.

4.8 Exercises.

4.1 Consider the following variant of the aircraft speed dynamics (Example 1.1). Let v be the deviation from the speed of lowest drag and let T be the thrust of the jet engine. If the control signal u is the desired thrust, then the following equations hold under certain operating conditions.

$$\begin{aligned}\dot{T} &= -T + u \\ \dot{v} &= -v^2 + T \\ y &= v\end{aligned}$$

where we have assumed that the speed is measured.
a. Construct an observer based on differentiation. **b.** Construct an observer with linear error dynamics (if possible).

4.2 Consider the system

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_2 x_3 \\ \dot{x}_3 &= x_2 \\ y_1 &= x_1 \\ y_2 &= x_3\end{aligned}$$

Construct an observer with linear error dynamics (if possible).

4.3 A ship that moves with a constant speed in a straight line is observed with a radar that measures distance only. Let p_x and p_y be the position of the ship in rectangular coordinates, v its speed through water, and θ its heading angle. If y is the radar measurement, the dynamical equations are

$$\begin{aligned}\dot{p}_x &= v\cos\theta \\ \dot{p}_y &= v\sin\theta \\ \dot{v} &= 0 \\ \dot{\theta} &= 0 \\ y &= \sqrt{p_x^2 + p_y^2}\end{aligned}$$

a. Construct an observer based on differentiation. **b.** Construct an observer with linear error dynamics (if possible). **c.** Let there be a second measurement $y_2 = \theta$. Construct an observer with linear error dynamics (if possible).

4.4 The roll dynamics of a certain missile is described by

$$\begin{aligned}\dot{\phi} &= \omega \\ \dot{\omega} &= -\beta \sin 4\phi - \alpha\omega + d \\ \dot{d} &= -d + u\end{aligned}$$

where ϕ is the roll angle, ω the angular velocity, d the aileron angle and u the control input. The sin-term is caused by aerodynamic vortices. The roll angle ϕ is observed.
a. Construct an observer based on differentiation. **b.** Construct an observer with linear error dynamics (if possible).

4.5 Consider Example 4.4. If the acceleration is a constant unknown, the system equations are

$$\begin{aligned}\dot{x} &= v \\ \dot{v} &= a - v^2 \\ \dot{a} &= 0 \\ y &= x\end{aligned}$$

Construct an observer with linear error dynamics (if possible).

4.6 Let a system be described in obsevable form by

$$\begin{aligned}\dot{x} &= f(x) + g(x)u \\ y &= h(z)\end{aligned}$$

and in observer form by

$$\begin{aligned}\dot{\xi} &= b(C\xi) + A\xi + \beta(\xi)u \\ \tilde{y} &= C\xi\end{aligned}$$

If we want an observer with linear error dynamics, then $\beta(\xi)$ must be a function of $C\xi = \xi_1$ only. What conditions must $g(x)$ then satisfy?

Chapter 5

Stability and Lyapunov theory.

Stability is probably the most important property of a control system. If a system is unstable , then it becomes meaningless to discuss properties like stationary errors, sensitivity or disturbance rejection. For nonlinear systems the definition of stability requires care. For instance it is quite possible that a system converges to a desired equilibrium after a small disturbance, but breaks into oscillation after a large one. The best tool for investigating stability related questions for nonlinear systems is Lyapunov theory. In fact, it is essentially the only tool.

5.1 Lyapunov functions.

Consider a nonlinear system described by

$$\dot{x} = f(x) \tag{5.1}$$

where x is an n-vector and f is a continuously differentiable function. From Chapter 2 it follows that for every starting point there is a unique solution (at least on a small time interval). Let $\pi(t,x)$ denote the solution of (5.1) at time t ,when the initial value at $t = 0$ is x, (i.e. $\pi(0,x) = x$).

To study the behavior of solutions it is convenient to introduce a scalar function which is analogous to the potential of physics.

Definition 5.1 A continuously differentiable scalar function V is called a *Lyapunov function* of (5.1) on G if $V_x(x)f(x) \le 0$ for $x \in G$. (G is any open subset of R^n).

From the definition the following fundamental property of Lyapunov functions immediately follows.

Proposition 5.1 If V is a Lyapunov function on G , then for $t \geq 0$ and as long as a trajectory remains in G, one has

$$V(\pi(t,x)) \leq V(x) \tag{5.2}$$

Proof.

$$V(\pi(t,x)) - V(x) = \int_0^t \frac{d}{d\tau} V(\pi(\tau,x))d\tau = \int_0^t V_x(\pi(\tau,x))f(\pi(\tau,x))d\tau \leq 0$$

■

Now consider sets of the form

$$B_d = \{x : V(x) \leq d\}, \quad \text{where } d > 0 \tag{5.3}$$

What one tries to do is to find a Lyapunov function such that for some $d > 0$ the set B_d lies in G. If this is possible one knows immediately that solutions starting in B_d never leave it.

Proposition 5.2 If V is a Lyapunov function on G and for some $d > 0$ the set B_d lies in G, then for all $x \in B_d$ one has $\pi(t,x) \in B_d$, for all $t > 0$.

Proof. Follows immediately from Proposition 5.1. ■

With good luck one might find that B_d is bounded. Then the following stability result follows.

Proposition 5.3 Let the conditions of Proposition 5.2 hold and assume that B_d is bounded. Then all solutions of (5.1) starting in B_d exist for all t and are bounded.

Proof. Follows from the previous proposition and Theorem 2.2. ■

One way of ensuring that the sets B_d are bounded is to use a function V which is *radially unbounded* i.e. satisfies

$$V(x) \to \infty \text{ as } |x| \to \infty$$

If it is possible to use Proposition 5.3 to show that solutions are bounded, then the next step is to find out where they go as t goes to infinity. This is formalized by the concept of a limit set

Definition 5.2 Let the conditions of Proposition 5.2 hold and let x be a point in B_d. A point y in B_d belongs to the *limit set* $\Omega(x)$ if there is a sequence $\{t_i\}$ such that $t_i \to \infty$ and $\pi(t_i,x) \to y$. In other words, solutions starting at x come arbitrarily close to the set $\Omega(x)$.

When a suitable Lyapunov function is known, it becomes possible to find out where the limit set is, as shown by the following theorem.

Theorem 5.1 (Lyapunov, LaSalle) Let the assumptions of Proposition 5.3 be satisfied and define

$$E = \{x \in B_d : V_x(x)f(x) = 0\}$$

Let M be the largest set in E with the property that solutions of (5.1) starting in M remain in M for all $t > 0$. (Such a set is called an *invariant set*). Then for each $x \in B_d$ there exists a c such that $\Omega(x) \subset M \cap V^{-1}(c)$.

Proof. Let x_0 be an arbitrary point in B_d. V is continuous on the compact set B_d and hence bounded below on that set. Since it is also decreasing along solutions of (5.1), (Proposition 5.1) , there exists a limit : $\lim_{t\to\infty} V(\pi(t, x_0)) = c$. Let y be an arbitrary point in $\Omega(x_0)$. Then there exists a sequence $\{t_i\}$, $t_i \to \infty$, such that $\pi(t_i, x_0) \to y$. Since V is continuous, it follows that $V(y) = c$. Hence $\Omega(x_0) \subset V^{-1}(c)$. From the relationship

$$\pi(t, y) = \lim_{t_i\to\infty} \pi(t, \pi(t_i, x_0)) = \lim_{t_i\to\infty} \pi(t + t_i, x_0)$$

it follows that $\pi(t, y) \in \Omega(x_0)$ for all $t > 0$. Then $c = V(\pi(t, y))$ so that

$$0 = \frac{d}{dt} V(\pi(t, y)) = V_x(\pi(t, y))f(\pi(t, y))$$

which shows that $\Omega(x_0) \subset M$. ■

The following example shows the kind of situation that can occur.

Example 5.1 Let the differential equation be

$$\begin{array}{rcl} \dot{x}_1 & = & x_1 + x_2 - x_1^3 - x_1x_2^2 \\ \dot{x}_2 & = & -x_1 + x_2 - x_1^2x_2 - x_2^3 \end{array} \tag{5.4}$$

Defining $r^2 = x_1^2 + x_2^2$ and the function $V(x) = r^4 - 2r^2$ gives

$$V_x(x)f(x) = -4r^2(r^2 - 1)^2 \leq 0$$

so V is a Lyapunov function everywhere. Also the set B_d is bounded for any $d > 0$. Since

$$M = E = \{(0, 0)\} \cup \{x_1^2 + x_2^2 = 1\}$$

it follows that the limit set is either the origin or the unit circle. ■

The situation that is of most interest in applications is described by the following Corollary.

Corollary 5.1 Assume that the conditions of Theorem 5.1 are satisfied and that M consists of a single point $\{x_m\}$. Then all solutions starting in B_d converge to x_m, which is necessarily a singular point of (5.1).

Proof. Since $\Omega(x_0) \subset \{x_m\}$,it follows that all solutions converge to x_m. Since solutions starting in $\{x_m\}$ remain in $\{x_m\}$, it follows that $f(x_m) = 0$. ■

The corollary can be illustrated by a harmonic oscillator with nonlinear damping.

Example 5.2 Consider the system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - x_2^3 \end{aligned} \tag{5.5}$$

and the function $V = x_1^2 + x_2^2$. Differentiation gives $V_x f = -2x_2^4 \leq 0$ so V is a Lyapunov function where E is the x_1-axis . Since a solution remaining in E must satisfy $x_2 \equiv 0$, it follows that also $x_1 = 0$, so $M = \{(0,0)\}$. Since B_d is bounded no matter how large d is, all solutions of the differential equation approach the origin. ■

5.2 Non-autonomous systems

Often one encounters nonlinear systems where the right hand side depends explicitly on the time variable, so called *non-autonomous systems*:

$$\dot{x} = f(t, x) \tag{5.6}$$

We will look at the stability of equilibrium points for such a system. To simplify things, we assume that the equilibrium is $x = 0$. The following definitions are straightforward extensions of those that are used for time-invariant systems.

Definition 5.3 The system (5.6) with the equilibrium $x = 0$ is

- *uniformly stable* if for each $\epsilon > 0$ there is a $\delta(\epsilon) > 0$, independent of t_o, such that
$$|x(t_o)| < \delta(\epsilon) \ \Rightarrow \ |x(t)| < \epsilon, \text{ all } t \geq t_o$$
- *uniformly asymptotically stable* if it is uniformly stable, and there exists a $\delta_o > 0$ such that
$$|x(t_o)| < \delta_o \ \Rightarrow \ |x(t)| \to 0, t \to \infty$$
and the convergence is uniform in t_o.
- *globally uniformly asymptotically stable* if the definition above holds for any $\delta_o > 0$.

Stability for non-autonomous systems can be proved using Lyapunov functions. Since the differential equation is time-varying it is however natural to allow time-varying Lyapunov functions. To handle the time-variability we will assume that the Lyapunov function is shut in between two positive definite functions. A function W is called *positive definite* if it satisfies $W(0) = 0$, $W(x) > 0$ for $x \neq 0$.

Theorem 5.2 Let $x = 0$ be an equilibrium of (5.6). Assume that there is an open connected set around the origin where the following relations hold for all t:

$$W_1(x) \leq V(t, x) \leq W_2(x) \tag{5.7}$$

$$\frac{d}{dt} V(t, x) = V_t + V_x f(t, x) \leq -W_3(x) \tag{5.8}$$

Here V is a continuously differentiable function, and the W_i are continuous positive definite functions. Then $x = 0$ is uniformly asymptotically stable.

Proof. Let the set where the conditions of the theorem hold be Ω. Define $\Omega_r = \{x : |x| \leq r\}$ and take an r such that $\Omega_r \subset \Omega$. Take a ρ such that

$$\rho < \min_{|x|=r} W_1(x)$$

and define the sets

$$\Omega_V = \{x \in \Omega_r : V(t,x) \leq \rho\}, \;\; \Omega_i = \{x \in \Omega_r : W_i(,x) \leq \rho\}, \;\; i = 1,2$$

Note that Ω_V is time dependent. From the construction it follows that

$$\Omega_2 \subset \Omega_V \subset \Omega_1 \subset \Omega_r$$

Now assume $x(t_o) \in \Omega_2$. Then $x(t_o) \in \Omega_V$ and consequently $V(t_o, x(t_o)) \leq \rho$. Since V has a negative derivative, it is decreasing and consequently $V(t, x(t)) \leq \rho$ holds for all t. It follows that $x(t) \in \Omega_1 \subset \Omega_r$ for all $t \geq t_o$. Since this holds for arbitrarily (small enough) $r > 0$, uniform stability is clear. To show that solutions actually converge to the origin, we use the fact that any continuous positive definite function $W(x)$ can be bounded from below and above as

$$\alpha_m(|x|) \leq W(x) \leq \alpha_M(|x|), \;\; x \in \Omega_r$$

where α_m, α_M are strictly increasing functions with $\alpha_m(0) = \alpha_M(0) = 0$. We can thus write

$$\alpha_1(|x|) \leq V(t,x) \leq \alpha_2(|x|), \quad \dot{V} \leq -\alpha_3(|x|)$$

where the α_i are strictly increasing functions with $\alpha_i(0) = 0$. We now get

$$\dot{V} \leq -\alpha_3(|x|) \leq -\alpha_3(\alpha_2^{-1}(V))$$

V is thus bounded above by the solution to the differential equation $\dot{z} = -\alpha_3(\alpha_2^{-1}(z))$ whose solution must converge to the origin. It follows that V decreases to zero. From $\alpha_1(|x|) \leq V(t,x)$ it follows that x converges to the origin. ■

There is a corollary about global stability.

Corollary 5.2 Let the assumptions of Theorem 5.2 be satisfied for all x and assume that W_1 is radially unbounded. Then $x = 0$ is globally uniformly asymptotically stable.

Proof. If W_1 is radially unbounded so is W_2. It follows that r in the proof can be taken arbitrarily large. ■

5.3 Construction of Lyapunov functions

In most situations one has a system with a known critical point, i.e. the point x_m of Corollary 5.1 is known, and the problem is to find a Lyapunov function V

which proves that solutions of (5.1) converge to x_m. To get an idea of the problems involved, it is instructive to investigate how this might be done for linear systems. To simplify the notation we assume that x_m is the origin. Consider the n-th order linear system

$$\dot{x} = Ax \tag{5.9}$$

where A has all its eigenvalues strictly in the left half plane and try to find a positive definite Lyapunov function, which is a degree m homogeneous polynomial in x. We can then write it in the form

$$V(x) = \sum v_i x_1^{k_{i1}} x_2^{k_{i2}} \cdots x_n^{k_{in}} \tag{5.10}$$

where the sum is taken over all k_{ij} such that $k_{i1} + k_{i2} + \cdots + k_{in} = m$. Now let $W(x)$ be a given positive definite homogeneous polynomial of degree m. Then it is natural to try to solve

$$V_x(x)Ax = -W(x) \tag{5.11}$$

If (5.11) is solved, then the conditions of Corollary 5.1 are automatically satisfied with $x_m = 0$ and B_d can be made as large as desired.

Theorem 5.3 (Lyapunov) If A has all its eigenvalues strictly in the left half plane, then (5.11) can always be solved with respect to V.

Proof. V_x consists of polynomials of degree $m-1$, since the differentiation lowers the degree by one. Multiplying by Ax gives a degree m homogeneous polynomial again. By identifying the coefficients of the monomials in the left and right hand sides of (5.11) we get a set of linear equations for the coefficients of V. By using e. g. Kronecker products one can show that this system is nonsingular whenever A has all its eigenvalues strictly in the left half plane. ■

Remark 5.1 When V and W are quadratic they are usually written in the form

$$V(x) = x^T Sx, \quad W(x) = x^T Qx$$

and equation (5.11) takes the form

$$A^T S + SA = -Q$$

called Lyapunov's equation.

Example 5.3 Let the system be given by

$$\dot{x} = \begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix}$$

and let the desired W be $x_1^4 + x_2^4$. Writing V in the form

$$V(x) = v_1 x_1^4 + v_2 x_1^3 x_2 + v_3 x_1^2 x_2^2 + v_4 x_1 x_2^3 + v_5 x_2^4$$

the left hand side of (5.11) becomes

$$-4v_1 x_1^4 + (4v_1 - 5v_2)x_1^3 x_2 + 3(v_2 - 2v_3)x_1^2 x_2^2 + (2v_3 - 7v_4)x_1 x_2^3 + (v_4 - 8v_5)x_2^4$$

An identification of coefficients gives

$$v_1 = 1/4, \quad v_2 = 1/5, \quad v_3 = 1/10, \quad v_4 = 2/70, \quad v_5 = 9/70$$

■

Now consider a general system

$$\dot{x} = f(x)$$

Assume that f is given by a series expansion

$$f(x) = f^{(1)}(x) + f^{(2)}(x) + f^{(3)}(x) + \cdots$$

where $f^{(k)}(x)$ is a homogeneous polynomial of degree k. Consider the problem of finding a Lyapunov function V such that

$$V_x(x) f(x) = -W(x) \tag{5.12}$$

where W is a given positive definite function. Assume that V and W can be written as expansions

$$V(x) = V^{(2)}(x) + V^{(3)}(x) + \cdots$$
$$W(x) = W^{(2)}(x) + W^{(3)}(x) + \cdots$$

Substituting into (5.12) and identifying coefficients gives

$$\begin{array}{rcl} V_x^{(2)}(x) f^{(1)}(x) & = & -W^{(2)}(x) \\ V_x^{(3)}(x) f^{(1)}(x) & = & -W^{(3)}(x) - V^{(2)}(x) f^{(2)}(x) \\ V_x^{(4)}(x) f^{(1)}(x) & = & -W^{(2)}(x) - V^{(3)}(x) f^{(2)}(x) - V^{(2)}(x) f^{(3)}(x) \\ & \vdots & \end{array} \tag{5.13}$$

Here the linear part, $f^{(1)}(x)$ can be written in the more familiar form

$$f^{(1)}(x) = Ax \tag{5.14}$$

where A is an n by n matrix. An examination of this system of equations leads to the following result.

Theorem 5.4 If f and W are given real analytic functions and if A of (5.14) has all its eigenvalues strictly in the left half plane, then it is possible to find a real analytic function V which solves (5.12). The coefficients of a series expansion for V can be computed from (5.13).

Proof. Since A has all its eigenvalues strictly in the left half plane, we can use Theorem 5.3, which tells us that any equation in (5.13) can be solved (if its right hand side is known). One can then solve the first equation of (5.13) to get $V^{(2)}$. Substituting $V^{(2)}$ into the right hand side of the second equation, one can solve for $V^{(3)}$. If $V^{(2)}$ and $V^{(3)}$ are substituted into the third equation, $V^{(4)}$ can be computed, and so on. The resulting series can be shown to be convergent in some neighborhood of the origin, so that V is well defined there. ■

5.4 Lyapunov functions and frequency domain criteria.

A very common structure for nonlinear systems is the one given in Figure 5.1. The system can be separated into a linear block with transfer function $G(s)$ and a nonlinear one, arranged in a feedback loop. We will consider the case of a static but possibly time-varying nonlinearity

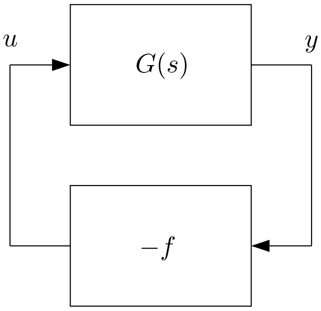


Figure 5.1: Nonlinear system separated into a linear and a nonlinear part.

$$u = -f(t, y) \tag{5.15}$$

The stability of such a configuration is a classical problem in control and its study goes back at least to the second world war and work by Lure. We assume that the linear part has m inputs and m outputs and that it has a minimal n-dimensional realization

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad G(s) = C(sI - A)^{-1}B \tag{5.16}$$

Lure's work was inspired by applications where the nonlinearity was not precisely known, so he only assumed that it satisfied certain inequalities. We will assume that the inequality is of the form

$$f(t, y)^T(f(t, y) - Ky) \leq 0, \quad \text{all } t, \text{ all } y \tag{5.17}$$

where K is a positive definite matrix. In the scalar case, $m = 1$, this inequality has a simple interpretation, shown in figure 5.2. The nonlinear function is bound by the horizontal axis and a line with slope K.

We will try to construct a quadratic Lyapunov function

$$V(x) = x^T P x \tag{5.18}$$

where P is a positive definite matrix. Computing the time derivative we get

$$\dot{V} = (Ax - Bf)^T P x + x^T P(Ax - Bf) \tag{5.19}$$

Subtracting the negative quantity given by the left hand side of (5.17) we get

$$\dot{V} \leq x^T(AP + PA)x - 2x^T PBf - 2f^T(f - Ky) \tag{5.20}$$

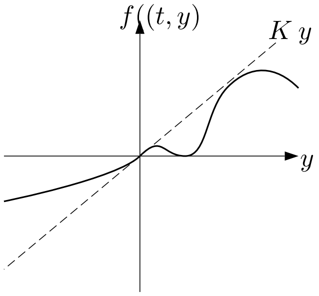


Figure 5.2: Bounds on the nonlinearity in the scalar case.

Completing the square we get

$$\dot{V} \le x^T(AP + PA + L^T L)x - (\sqrt{2}f - Lx)^T(\sqrt{2}f - Lx) \tag{5.21}$$

where $L = (1/\sqrt{2})(KC - B^T P)$. Suppose now that we can find a P such that

$$A^T P + PA = -L^T L - \epsilon P \tag{5.22}$$

$$PB = C^T K - \sqrt{2}L^T \tag{5.23}$$

Then we get

$$\dot{V} \le -\epsilon x^T P x - (\sqrt{2}f - Lx)^T(\sqrt{2}f - Lx) \tag{5.24}$$

which is negative definite and we can use Theorem 5.2 to prove stability. The solvability of (5.22),(5.23) is characterized by a classical result.

Lemma 5.1 The Kalman-Yakubovich-Popov lemma. Assume that A has eigenvalues with strictly negative real parts and that A, C is observable, A, B controllable. Then the matrix equations

$$\begin{aligned} A^T P + PA &= -L^T L - \epsilon P \\ PB &= C^T - L^T W \\ W^T W &= D + D^T \end{aligned}$$

can be solved for L, W, $P > 0$, $\epsilon > 0$ if and only if the transfer function

$$C(sI - A)^{-1}B + D$$

is strictly positive real.

Proof. See the Appendix. ■

In this context we use the following definition of positive real.

Definition 5.4 A square rational transfer function matrix G is called positive real if

- all elements of $G(s)$ are analytic for Re $s > 0$,

- any pure imaginary pole is a simple pole with positive semidefinite residue matrix,
- for all real ω the matrix $G(i\omega) + G^T(-i\omega)$ is positive semidefinite.

$G(s)$ is called strictly positive real if $G(s - \epsilon)$ is positive real for some $\epsilon > 0$.

It follows that an input-output stable SISO system is positive real when its Nyquist curve lies in the closed right half plane.

We are now ready to state the basic stability result.

Theorem 5.5 Let the system (5.16) have all its poles strictly in the left half plane and let the nonlinear feedback (5.15) satisfy the inequality (5.17). Then the closed loop system is globally asymptotically stable if $I + KG(s)$ is strictly positive real.

Proof. If $I + KG(s)$ is strictly positive real it follows from the Kalman-Yakubovich-Popov lemma (with $D = I$ and C replaced by KC) that (5.22), (5.23) is solvable. $V = x^T Px$ is then a Lyapunov function satisfying the conditions of Theorem 5.2. ■

The theorem can be reformulated using a standard trick called pole shifting, see Figure 5.3. Adding and subtracting the linear block K_1 does not alter the loop.

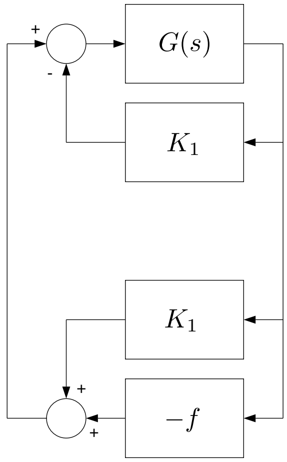


Figure 5.3: Pole shifting by adding and subtracting a linear block.

However, the new diagram can be interpreted in a different way, as a linear system

$$\tilde{G} = G(I + K_1 G)^{-1}$$

with the nonlinear feedback

$$\tilde{f}(t, y) = f(t, y) - K_1 y$$

Suppose the nonlinearity satisfies the condition

$$(f(t,y) - K_1 y)^T (f(t,y) - K_2 y) \leq 0, \quad \text{all } t, \text{ all } y \tag{5.25}$$

then the modified nonlinearity satisfies

$$\tilde{f}(t,y)(\tilde{f}(t,y) - (K_2 - K_1)y) \leq 0$$

and we get directly the corollary.

Corollary 5.3 Let the system (5.16) have a nonlinear feedback (5.15) satisfying the inequality (5.25), with $K_2 - K_1$ positive definite. Then the closed loop system is globally uniformly asymptotically stable if

$$\tilde{G} = G(I + K_1 G)^{-1}$$

has all its poles strictly in the left half plane and

$$(I + K_2 G(s)(I + K_1 G(s))^{-1}$$

is strictly positive real.

Proof. As shown above the nonlinearity $\tilde{f}$ satisfies (5.17) with $K = K_2 - K_1$. From Theorem 5.5 we have to check the strict positive realness of

$$\begin{aligned} I + (K_2 - K_1)\tilde{G}(s) = I + (K_2 - K_1)G(s)(I + K_1 G(s))^{-1} = \\ = (I + K_2 G(s))(I + K_1 G(s))^{-1} \end{aligned}$$

■

For the scalar case $m = 1$ this corollary can be formulated geometrically in a well-known way.

Corollary 5.4 The Circle Criterion. Let u and y of the linear system (5.16) be scalars. Assume that G has no poles in the right half plane and let the nonlinearity satisfy

$$k_1 y^2 \leq y f(t,y) \leq k_2 y^2, \qquad \text{all } t, \text{ all } y \tag{5.26}$$

Then a sufficient condition for the closed loop system to be globally asymptotically stable is that the Nyquist curve of G does not encircle or enter the circle whose diameter lies on the real axis between $-1/k_1$ and $-1/k_2$.

The geometric interpretation of the circle criterion is given in Figure 5.4.

Proof. The condition $k_1 y^2 \leq y f(t,y) \leq k_2 y^2$ can also be rewritten $(f - k_1 y)(f - k_2 y) \leq 0$. The stability condition is from Corollary 5.3 that $G/(1 + k_1 G)$ has all its poles in the left half plane, and that

$$\frac{1 + k_2 G}{1 + k_1 G}$$

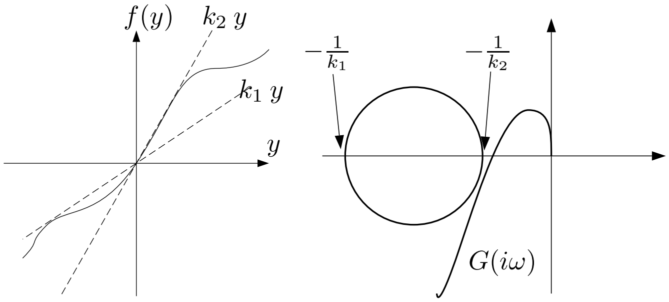


Figure 5.4: The circle criterion. Bounds on the nonlinearity to the left and bounds on the Nyquist curve to the right.

is positive real. Since the rational function

$$\frac{1+k_2 z}{1+k_1 z}$$

maps the interior of the circle going through $-1/k_1$, $-1/k_2$ onto the left half plane, the positive realness condition is equivalent to the condition that the Nyquist curve does not enter the circle. The stability of $G/(1+k_1 G)$ follows if the Nyquist curve does not enclose the circle. ■

There is a variation on the results above that uses a slightly different Lyapunov function for the time-invariant case, namely

$$V = x^T P x + 2\eta \int_0^y f^T(\sigma) K \, d\sigma \tag{5.27}$$

where $\eta \geq 0$ is a free parameter. It is assumed that $f^T K$ is the gradient of a scalar function so that the integral is path-independent (this is automatically satisfied in the scalar case). To get a reasonable V, it is assumed that

$$\int_0^y f^T(\sigma) K \, d\sigma \geq 0, \text{ all } y \tag{5.28}$$

The nonlinearity is assumed to satisfy

$$f^T(y)(f(y) - Ky) \leq 0, \text{ all } y \tag{5.29}$$

Differentiating (5.27) gives

$$\dot{V} = x^T(A^T P + PA)x - 2x^T PBf + 2\eta f^T KC(Ax - Bf)$$

Adding the nonnegative quantity $-2f^T(f - Ky)$ gives

$$\dot{V} \leq x^T(A^T P + PA)x - f^T(2I + \eta KCB + \eta B^T C^T K)f + \\ + 2x^T(-PB + \eta A^T C^T K + C^T K)f$$

If we factorize according to

$$2I + \eta KCB + \eta B^T C^T K = W^T W \tag{5.30}$$

then we can complete the squares to get

$$\dot{V} \leq x^T(A^TP + PA + L^TL)x - (Wf - Lx)^T(Wf - Lx)$$

provided L satisfies

$$PB = C^TK + \eta A^TC^TK - L^TW \tag{5.31}$$

If we finally can satisfy

$$A^TP + PA = -L^TL - \epsilon P \tag{5.32}$$

for some $\epsilon > 0$, then we have

$$\dot{V} \leq -\epsilon x^TPx \tag{5.33}$$

The equations (5.32), (5.31), (5.30) are of the form occuring in the Kalman-Yakubovich-Popov lemma and we get the following result.

Theorem 5.6 The multivariable Popov criterion. Let the system (5.16) have all its poles strictly in the left half plane and let the nonlinear time invariant feedback $u = -f(y)$ satisfy the inequality (5.29). Also assume that f^TK is the gradient of a scalar function and that (5.28) is satisfied. Then the closed loop system is globally asymptotically stable if there is an $\eta \geq 0$ such that $-1/\eta$ is not an eigenvalue of A and

$$I + (1 + \eta s)KG(s)$$

is strictly positive real.

Proof. Applying the Kalman-Yakubovich-Popov lemma with C^T replaced by $C^TK + \eta A^TC^TK$ and $D = I + \eta KCB$ shows that we have to test for positive realness of

$$\begin{aligned}
&I + \eta KCB + (KC + \eta KCA)(sI - A)^{-1}B = \\
&\qquad I + KC(\eta I + (sI - A)^{-1} + \eta A(sI - A)^{-1})B = \\
&\qquad I + KC(\eta(sI - A) + I + \eta A)(sI - A)^{-1}B = \\
&\qquad\qquad I + (1 + \eta s)KG(s)
\end{aligned}$$

The condition that $-1/\eta$ is not an eigenvalue of A guarantees that the pair $C+\eta CA, A$ is observable if C, A is (we have assumed in the Kalman-Yakubovich-Popov lemma that the system is controllable and observable). ■

There is a simple geometric interpretation in the SISO case.

Corollary 5.5 The classical Popov criterion. Let the conditions of Theorem 5.6 be satisfied for a single-input-single-output system. Then a sufficient condition for global asymptotic stability is that

$$\frac{1}{K} + \operatorname{Re} G(i\omega) - \eta\omega \operatorname{Im} G(i\omega) > 0, \text{ all } \omega \tag{5.34}$$

Graphically this means that the so called Popov plot ($\omega\mathrm{Im}G(i\omega)$ versus $\mathrm{Re}G(i\omega)$) has to lie to the right of a line through $-1/K$ with slope $1/\eta$ for some value of η, see Figure 5.5.

Proof. In the scalar case the positive definiteness condition is that

$$\mathrm{Re}\,(1 + (1 + \eta i\omega)KG(i\omega)) > 0$$

Taking the real part of the expression within parentheses gives the result. ■

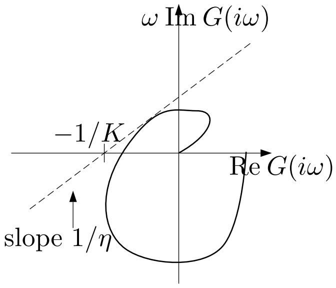


Figure 5.5: The classical Popov criterion. If the Popov plot lies to the right of the line, global asymptotic stability is guaranteed.

Example 5.4 Consider the linear system

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

in feedback with a saturation nonlinearity

$$f(y) = \begin{cases} Ky & |y| \le 1/K \\ 1 & y > 1/K \\ -1 & y < -1/K \end{cases}$$

In the left part of Figure 5.6 the Nyquist curve is shown together with a vertical line through $-1/9$. It is clear that $K \approx 9$ is the highest value for which the circle criterion guarantees stability in this case. In the right hand part of the figure the Popov curve is shown together with a line through $-1/100$, so the Popov criterion shows stability for $K = 100$. In fact it is clear that the line can intersect the real axis arbitrarily close to the origin, so the Popov criterion proves stability for any $K > 0$ (which is the stability result what one intuitively would guess). ■

5.5 Lyapunov functions and input output stability

Now consider a dynamic system

$$\dot{x} = f(x) + g(x)u, \quad y = h(x) \tag{5.35}$$

We make the following assumptions

$$f(0) = 0, \quad h(0) = 0, \quad |g(x)| \le c_1, \quad |h(x)| \le c_2|x| \tag{5.36}$$

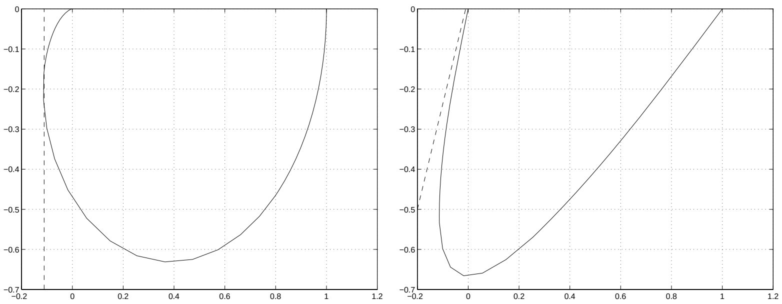


Figure 5.6: Comparison of circle and Popov criterion.

So far we have only studied the uncontrolled system

$$\dot{x} = f(x) \tag{5.37}$$

One might wonder if asymptotic stability of (5.37) also implies a nice input output behavior for (5.35). For linear systems it is well known that such a relationship exists. We will prove a similar result for the nonlinear case. To do that we assume that we have found a positive definite Lyapunov function V such that

$$V(0) = 0, \quad V_x(x)f(x) \leq -c_3|x|^2, \quad |V_x(x)| \leq c_4|x| \tag{5.38}$$

Then we can show the following

Theorem 5.7 Suppose the system (5.35) satisfies (5.36) and that a Lyapunov function exists with the properties (5.38). Then there exists a positive constant k and a function $\gamma(x)$ with $\gamma(0) = 0$, such that for any $T > 0$ and any input u

$$\int_0^T y^T(t)y(t)dt \leq k^2 \int_0^T u^T(t)u(t)dt + \gamma(x_0) \tag{5.39}$$

where x_0 is the initial state.

Proof. Using (5.36) and (5.38) we get

$$0 \leq V(x(T)) = \int_0^T \dot{V}(x(t))\,dt + V(x_0) = \int_0^T V_x(x(t))\,(f(x(t)) + g(x(t))u(t))\;dt$$

$$+V(x_0) \leq \int_0^T (-c_3|x(t)|^2 + c_1c_4|x(t)||u(t)|\;dt + V(x_0) =$$

$$-\frac{c_3}{2}\int_0^T |x(t)|^2\;dt - \int_0^T \left(\frac{c_3}{2}\left(|x(t)| - \frac{c_1c_4}{c_3}|u(t)|\right)^2 + \frac{c_1^2c_4^2}{2c_3}|u(t)|^2\right)\;dt + V(x_0)$$

showing that

$$\frac{c_3}{2}\int_0^T |x(t)|^2\;dt \leq \frac{c_1^2c_4^2}{2c_3}\int_0^T |u(t)|^2\;dt + V(x_0)$$

Since

$$|y(t)| \leq c_2|x(t)|$$

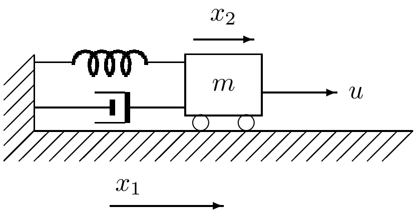
from (5.36), it follows that (5.39) is satisfied with

$$k = \frac{c_1c_2c_4}{c_3}, \quad \gamma(x_0) = \frac{2c_2^2}{c_3V(x_0)}$$

■

Remark 5.2 The inequality ((5.39) can often be given an energy interpretation: the energy in the output is less than some multiple of the energy in the input plus the stored energy.

Example 5.5 Consider the mechanical system



A mass m moving with velocity x_2, is pulled by an external force u. The mass is connected to a nonlinear spring giving the force

$$F = k_1x_1 + k_2x_1^3$$

and to a linear damper giving the force

$$F = bx_2$$

We consider the velocity x_2 to be the output. The equations of motion are then

$$\begin{aligned} \dot{x}_1 &= x_2 \\ m\dot{x}_2 &= u - k_1x_1 - k_2x_1^3 - bx_2 \\ y &= x_2 \end{aligned} \tag{5.40}$$

Now consider an energy balance for the time interval $0 \leq t \leq T$. We have energy delivered to the system:

$$\int_0^T uy dt$$

increase in kinetic energy:

$$\frac{m}{2}\left(x_2(T)^2 - x_2(0)^2\right)$$

increase in potential energy of the spring:

$$\int_0^T \dot{x}_1(k_1x_1 + k_2x_1^3)dt = \frac{k_1}{2}x_1(T)^2 + \frac{k_2}{4}x_1(T)^4 - \frac{k_1}{2}x_1(0)^2 - \frac{k_2}{4}x_1(0)^4$$

energy lost in damper:

$$b\int_0^T y^2 dt$$

Summing the various contributions we get

$$\int_0^T \left(-by^2 + uy\right)\, dt + \gamma(x(0)) = \gamma(x(T)) \tag{5.41}$$

where

$$\gamma(x) = \frac{k_1}{2}x_1^2 + \frac{k_2}{4}x_1^4 + \frac{m}{2}x_2^2$$

Since γ is a nonnegative function, we can also write

$$\int_0^T \left(-by^2 + uy\right)\, dt + \gamma(x(0)) \geq 0 \tag{5.42}$$

Note that the function γ, representing the energy stored in the system, is also a Lyapunov function for the system with $u = 0$, since

$$\dot{\gamma} = \gamma(x)_x \dot{x} = -bx_2^2 \leq 0$$

■

Motivated by this example one can consider general systems with an equilibrium at the origin

$$\dot{x} = f(x,u), \quad y = h(x,u), \quad x(0) = x_0, \quad f(0,0) = 0, \quad h(0,0) = 0 \tag{5.43}$$

and define a *supply rate*

$$w(u,y) = y^TQy + 2y^TSu + u^TRu \tag{5.44}$$

which is a quadratic form in the input and output. Motivated by (5.42) we make the following definition.

Definition 5.5 A system (5.43) is called *dissipative* if there exists a function w of the form (5.44) and a function γ, $\gamma(0) = 0$ such that

$$\int_0^T w(u(t), y(t))\, dt + \gamma(x_0) \geq 0 \tag{5.45}$$

for all $T \geq 0$ and all control signals u, where y is the output of (5.43) for the initial condition $x(0) = x_0$.

We see that equation (5.39) implies that the system is dissipative ($w = -y^Ty + k^2u^Tu$). This case is much studied in the control literature and motivates the following definition.

Definition 5.6 A dissipative system with

$$w(u,y) = -y^Ty + k^2u^Tu$$

is said to have *finite gain*.

For an electric circuit where the u is the voltage and y the current (or vice versa), the product uy is the supplied power. This has motivated the following definition.

Definition 5.7 A dissipative system is called *passive* if w has the form

$$w(u, y) = u^T y$$

and *strictly passive* if w is of the form

$$w(u, y) = u^T y - \epsilon u^T u$$

for some $\epsilon > 0$.

Let us now define the *storage function* V as

$$V(x_0) = -\inf_{u,T} \int_0^T w(u(t), y(t))\, dt \tag{5.46}$$

where y is the solution of (5.43) with $x(0) = x_0$. The infimum is taken over all inputs u and all times $T \geq 0$. We now have the following generalization of (5.42).

Proposition 5.4 Assume that (5.45) holds for a system (5.43). Then the storage function V defined by (5.46) satisfies

$$0 \leq V(x) \leq \gamma(x), \quad V(0) = 0$$

and

$$V(x_0) + \int_0^T w(u(t), y(t))\, dt \geq V(x_T), \quad T \geq 0 \tag{5.47}$$

where y is the output corresponding to u with the initial condition x_0 and x_T is the state reached at $t = T$.

Proof. Since $T = 0$ is allowed in (5.46), it is clear that $V \geq 0$. From (5.45) it follows that $V \leq \gamma$. Now for any choice of u, $T \geq 0$, $T_1 \geq T$ we have

$$V(x_0) \geq -\int_0^T w(u, y)dt - \int_T^{T_1} w(u, y)dt$$

The inequality is then also true if the last integral is replaced by its infimum, giving

$$V(x_0) \geq -\int_0^T w(u, y)dt + V(x_T)$$

■

An immediate consequence of the proposition is that storage functions and Lyapunov functions are closely related.

Proposition 5.5 Consider a dissipative system where $Q \leq 0$. Then the storage function V (5.46) is a Lyapunov function for the uncontrolled system ($u = 0$).

Proof. Setting $u = 0$ in (5.47) gives

$$V(x_T) \leq V(x_0) + \int_0^T y^T Q y \, dt$$

which shows that V is decreasing along trajectories. ■

A problem is that the definition (5.46) does not give any guarantee that V is a smooth function. Suppose however that V is a continuously differentiable function. Then we can differentiate (5.47) to get

$$\dot{V} = V_x(c) f(x, u) \leq w(y, u) \tag{5.48}$$

Let us specialize to systems of the form

$$\dot{x} = f(x) + g(x)u, \quad y = h(x), \quad f(0) = 0, \quad h(0) = 0 \tag{5.49}$$

Equation (5.48) then becomes, for the case $S = 0$

$$V_x f \leq -V_x g u + h^T Q h + u^T R u \tag{5.50}$$

If $R > 0$ we can complete the square to get

$$\begin{aligned} V_x f \leq (u - \frac{1}{2}R^{-1}g^T V_x^T)^T R (u - \frac{1}{2}R^{-1}g^T V_x^T) + \\ + h^T Q h - \frac{1}{4} V_x g R^{-1} g^T V_x^T \end{aligned} \tag{5.51}$$

From this inequality we get

Theorem 5.8 Consider a system (5.49) which is dissipative with $S = 0$, $R > 0$. Assume that the storage function V of (5.46) is continuously differentiable. Then it satisfies the *Hamilton-Jacobi inequality.*

$$V_x f + \frac{1}{4} V_x g R^{-1} g^T V_x^T - h^T Q h \leq 0, \quad V(0) = 0 \tag{5.52}$$

Proof. Since (5.51) must be satisfied for any u, it is in particular true for $u = \frac{1}{2}R^{-1}g^T V_x^T$, which gives (5.52). ■

There is a converse theorem.

Theorem 5.9 Consider a system of the form (5.49). Suppose the Hamilton-Jacobi inequality (5.52) has a differentiable, nonnegative solution for matrices Q, R, $R > 0$. The the system is dissipative with supply rate

$$y^T Q y + u^T R u$$

Proof. Since $R > 0$ we can add a term

$$(u - \frac{1}{2}R^{-1}g^T V_x^T)^T R (u - \frac{1}{2}R^{-1}g^T V_x^T)$$

to the right hand side of (5.52) to get (5.51). Going through (5.50), (5.48) in reverse we get (5.47), from which (5.45) follows. ■

5.6 Exercises

5.1 Rotating machinery is often described by the equation

$$\dot{x}_1 = x_2 \tag{5.53}$$

$$\dot{x}_2 = -d\ x_2 - f(x_1) \tag{5.54}$$

where x_1 is the angle of rotation and x_2 is the angular velocity. The constant $d > 0$ represents some type of viscous friction or damping and f a restoring moment, $f(0) = 0$. Show that

$$V(x) = \frac{1}{2}x_2^2 + \int_0^{x_1} f(u)\,du$$

is a Lyapunov function and discuss under what conditions the state variables converge to the origin.

5.2 Consider the system described in the previous exercise. For an electric generator f has the form

$$f(x) = \sin x$$

Compute the Lyapunov function and sketch the area of the phase plane where solutions will converge to the origin.

5.3 Use the function

$$V(x) = \frac{x_1^2 + x_2^2 - x_1 x_2^3}{2(1 - x_1 x_2)}$$

to estimate the domain of attraction of the origin for

$$\dot{x}_1 = -x_1 + 2x_1^2 x_2 \tag{5.55}$$

$$\dot{x}_2 = -x_2 \tag{5.56}$$

5.4 Let f be a decoupled nonlinearity

$$f(t,y) = \begin{bmatrix} f_1(t,y_1) \\ \vdots \\ f_m(t,y_m) \end{bmatrix}$$

where each component satisfies a sector condition of the form shown in the left part of Figure 5.4, i. e.

$$\ell_i y_i^2 \le y_i f_i(t,y_i) \le k_i y_i^2$$

Rewrite these conditions in the form (5.25).

5.5 Suppose the nonlinearity f satisfies the condition

$$|f(t,y) - Ly| \le \gamma |y|$$

for all t and y, where the vertical bars denote the Euclidean norm. Rewrite this condition in the form (5.25).

5.6 A DC motor with axis resonances has the transfer function

$$G(s) = \frac{4}{s(s+1)(s^2+0.1s+4)}$$

It is controlled by a saturated P-controller. For what gain of the P-controller can stability be guaranteed?

5.7 The system

$$G(s) = \frac{10}{(s+1)(s+2)(s+10)}$$

is controlled by a P-controller with dead-zone:

$$u = -\begin{cases} K(y-1) & y > 1 \\ 0 & |y| \le 1 \\ K(y+1) & y < -1 \end{cases}$$

For what K-values can one guarantee stability?

5.8 The system

$$G(s) = \begin{bmatrix} \frac{2}{s+1} & \frac{3}{s+2} \\ \frac{1}{s+1} & \frac{1}{s+1} \end{bmatrix}$$

is controlled using

$$u_1 = -f_1(y_1), \quad u_2 = -f_2(y_2)$$

where

$$0 \le y_1 f_1(y_1) \le k_1 y_1^2, \quad 0 \le y_2 f_2(y_2) \le k_2 y_2^2$$

For what values of k_1 and k_2 can stability be guaranteed?

5.9 What is the Hamilton-Jacobi inequality for a linear system if one assumes a quadratic function V?

5.7 Appendix

Proof. The kalman-Yakubovich-Popov lemma. We will only give an outline of the proof.

Positive realness implies solvability of the matrix equations. Suppose that $G(s)$ is strictly positive real. Then there is some $\epsilon > 0$ such that $G(s-\epsilon/2)$ is positive real. Since A has all its poles strictly in the left half plane, it is possible to choose ϵ so that $A_\epsilon = A + \frac{\epsilon}{2}I$ also has all its poles there.

Step 1. Factorize

$$G(s-\epsilon/2) + G^T(-s-\epsilon/2) = V^T(-s)V(s) \tag{5.57}$$

where V is a transfer matrix with all its poles in the left half plane. For scalar transfer functions it is fairly easy to see that this can be done, since the left hand side is symmetric with respect to sign changes in s. Poles and zeros then are symmetricly placed with respect to the imaginary axis and $V(s)$, $V(-s)$ can be formed by picking poles and zeros from the correct half plane. The multivariable case is due to a classical result by Youla.

Step 2. Make a state space realization of the left hand side of (5.57) using the realization of G. This gives

$$\dot{x} = \begin{bmatrix} A_\epsilon & 0 \\ 0 & -A_\epsilon^T \end{bmatrix} x + \begin{bmatrix} B \\ C^T \end{bmatrix} u, \quad y = \begin{bmatrix} C & -B^T \end{bmatrix} x + (D+D^T)u \tag{5.58}$$

by parallell connection.

Step 3. Make a state space realization of the right hand side of (5.57) using a minimal realization of V, with matrices F, G, H, J. Since $V^T(-s)V(s)$ can be regarded as a series connection of this system and one with matrices $-F^T$, H^T, $-G^T$, J^T, we get

$$\dot{\tilde{z}} = \begin{bmatrix} F & 0 \\ H^T H & -F^T \end{bmatrix} z + \begin{bmatrix} G \\ H^T J \end{bmatrix} u, \quad y = \begin{bmatrix} J^T H & -G^T \end{bmatrix} \tilde{z} + J^T J u \tag{5.59}$$

Step 4. Make (5.59) into block diagonal form by using the transformation

$$z = \begin{bmatrix} I & 0 \\ K & I \end{bmatrix} \tilde{z}$$

where K satisfies

$$KF + F^T K + H^T H = 0 \tag{5.60}$$

This gives

$$\dot{z} = \begin{bmatrix} F & 0 \\ 0 & -F^T \end{bmatrix} z + \begin{bmatrix} G \\ H^T J + KG \end{bmatrix} u, \quad y = \begin{bmatrix} J^T H + G^T K & -G^T \end{bmatrix} \tilde{z} + J^T J u \tag{5.61}$$

Step 5. Show that (5.58) and (5.61) are minimal realizations of the left and right hand sides of (5.57). This is an exercise in the PBH test.

Step 6. Since (5.58) and (5.61) are both minimal realizations of the same transfer function there is a transformation

$$z = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} x$$

between them. Consider T_{12}. Transforming between (5.58) and (5.61) shows that

$$FT_{12} + T_{12}A_\epsilon^T = 0$$

Multiplying with e^{Ft} from the left and $e^{A_\epsilon^T t}$ from the right shows that

$$\frac{d}{dt}\left(e^{Ft}T_{12}e^{A_\epsilon^T t}\right) = 0$$

which gives

$$T_{12} = e^{Ft}T_{12}e^{A_\epsilon^T t}$$

Since the right hand side goes to zero, it follows that $T_{12} = 0$. In a similar manner we can show that $T_{21} = 0$. The transformation is thus block diagonal.

Step 7. Writing out the tranformation between (5.58) and (5.61) gives

$$\begin{aligned} J^T J &= D + D^T \\ F &= T_{11}A_\epsilon T_{11}^{-1} \\ G &= T_{11}B \\ J^T H + G^T K &= CT_{11}^{-1} \end{aligned}$$

Step 8. Take

$$\begin{aligned} W &= J \\ L &= HT_{11} \\ P &= T_{11}^T K T_{11} \end{aligned}$$

Plugging these values into the equations of Step 7 and using (5.60) shows that the matrix equations of the Kalman-Yakubovich-Popov lemma are satisfied.

Solvability of the matrix equations implies positive realness.

This is a straightforward but tedious calculation. Since it is not needed for the circle and Popov theorems, we omit it. ∎

.

Chapter 6

Lyapunov based controller design

Since stability is such a central issue for control, it is natural that many controller design methods use Lyapunov theory.

6.1 Control Lyapunov functions

A Lyapunov function is defined for a nonlinear system $\dot{x} = f(x)$ where the right hand side is given. However, when doing control design the system is $\dot{x} = f(x, u)$ where the control u is to be chosen. It is then natural to define a *control Lyapunov function* V to be a positive definite, radially unbounded function such that

$$x \neq 0 \;\Rightarrow\; V_x(x)f(x,u) < 0, \quad \text{for some } u \tag{6.1}$$

If a control Lyapunov function exists, then it is natural to try to pick a feedback control $u = k(x)$ such that $V_x(x)f(x, k(x)) < 0$, which would guarantee asymptotic stability. It is clear that such a choice can be made, but it is not clear that it can be done without discontinuities in $k(x)$. In particular, if the system has the form

$$\dot{x} = f(x) + g(x)u \tag{6.2}$$

then V being a control Lyapunov function implies that

$$V_x g = 0 \;\Rightarrow\; V_x f < 0 \tag{6.3}$$

It is then natural to use a control law of the form

$$u = -\phi(x)(V_x(x)g(x))^T \tag{6.4}$$

where ϕ is some positive scalar function.

Open loop stable systems

Consider a system of the form (6.2) which is open loop stable but not necessarily asymptotically stable. Suppose we know a function V such that

$$V_x f(x) \leq 0, \quad \text{all } x$$

If we have $V_x g \neq 0$ whenever $V_x f = 0$ for $x \neq 0$, then we have a control Lyapunov function. We could choose $\phi(x) = 1$ in (6.4) to achieve a smooth control law which gives a negative definite $\dot{V}$.

6.2 Backstepping

Backstepping is a method to extend a Lyapunov function and an associated control law from part of the system to all the system by moving backwards through the system. Typically one starts with a system

$$\dot{x} = f(x) + g(x)z \tag{6.5}$$

for which a control law $z = k(x)$ and a Lyapunov function V are known, that is

$$\dot{V} = V_x(f(x) + g(x)k(x)) = -W(x) \tag{6.6}$$

where W is some positive definite function. Now suppose this system is really part of a larger system, so that z is not directly available but instead the system description is

$$\dot{x} = f(x) + g(x)z \tag{6.7}$$

$$\dot{z} = a(x,z) + b(x,z)u \tag{6.8}$$

At this stage we assume for simplicity that z is a scalar. Rewriting the first equation gives

$$\begin{aligned} \dot{x} &= f(x) + g(x)k(x) + g(x)(z - k(x)) \\ \dot{z} &= a(x,z) + b(x,z)u \end{aligned}$$

Let us try to extend the given Lyapunov function by writing

$$V_e(x,z) = V(x) + \frac{1}{2}(z - k(x))^2 \tag{6.9}$$

The time derivative then becomes (omitting arguments)

$$\begin{aligned} \dot{V}_e = V_x(f + gk) + V_x g\,(z - k) + (z - k)(a + bu - k_x\,(f + gz)) = \\ -W + (V_x g + a + bu - k_x\,(f + gz))(z - k) \end{aligned}$$

We see that we can get

$$\dot{V}_e = -W - \gamma(z - k)^2$$

which is negative definite, if u is chosen as

$$u = \frac{1}{b}(k_x(f + gz) - a - V_x g - \gamma\,(z - k)) \tag{6.10}$$

We have thus defined a control law for the extended system such that V_e becomes a Lyapunov function.

Example 6.1 Consider a version of the mechanical system described in Example 6.5 where there is no damping and the spring force is $-x_1^3$. The system is then described by

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -x_1^3 + u\end{aligned}$$

if we neglect the disturbance w. By using the feedback

$$u = -x_1 - x_2$$

we get the same system as we used in Example 6.5. We thus know the Lyapunov function

$$V = \frac{3}{2}x_1^2 + x_1x_2 + x_2^2 + \frac{x_1^4}{2},$$

for that closed loop system. Now suppose that we apply the force through an actuator that has a certain time constant so that we have

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -x_1^3 + z\\ \dot{z} &= -z + u\end{aligned}$$

Using (6.10) we get

$$u = -(1+\gamma)x_1 - (3+\gamma)x_2 - \gamma z + x_1^3 \tag{6.11}$$

In Figure 6.1 a simulation is shown for the case $\gamma = 1$. As a comparison the response is also shown when only the linear part of the control law is used, i.e.

$$u = -(1+\gamma)x_1 - (3+\gamma)x_2 - \gamma z \tag{6.12}$$

It is clear that the nonlinear term is needed to get a good damping. ■

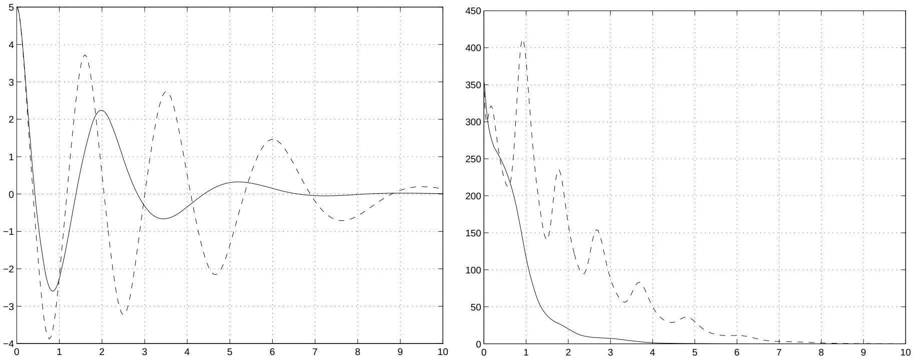


Figure 6.1: Response of mechanical system with backstepping controller (solid) compared to linear controller (dashed). The left diagram shows x_1 the right one shows the Lyapunov function.

The backstepping approach is easily extended to vector valued inputs. Suppose z and u in (6.7), (6.8) are vectors of length m. Then it is natural to replace (6.9) by

$$V_e(x,z) = V(x) + \frac{1}{2}(z - k(x))^T(z - k(x)) \tag{6.13}$$

and we get

$$\dot{V}_e = -W - \gamma(z-k)^T(z-k)$$

by taking

$$u = b^{-1}(k_x(f+gz) - a - (V_x g)^T - \gamma\,(z-k)) \tag{6.14}$$

It is also possible to handle the system structure

$$\begin{aligned}\dot{x} &= f(x,z)\\ \dot{z} &= a(x,z) + b(x,z)u\end{aligned}$$

provided the system $\dot{x} = f(x,u)$ has a stabilizing control law $u = k(x)$ and a Lyapunov function $V(x)$ with

$$V_x(x)f(x,k(x)) = -W(x), \quad W \text{ positive definite}$$

The idea is still to take the extended Lyapunov function

$$V_e(x,z) = V(x) + \frac{1}{2}(z - k(x))^2$$

and choose u to give

$$\dot{V}_e = -W(x) - (z - k(x))^2$$

The formulas become messier because the expression for u will contain a division with the factor $(z - k(x)$ which can not be explicitly cancelled.

Example 6.2 Consider the system

$$\begin{aligned}\dot{x}_1 &= \arctan x_2\\ \dot{x}_2 &= u\end{aligned}$$

Take to begin with the system

$$\dot{x}_1 = \arctan u$$

It can be stabilized with the control $u = -x_1$, as shown by taking $V = \frac{1}{2}x_1^2$, giving

$$\dot{V} = -x_1 \arctan x_1$$

The extended Lyapunov function then becomes

$$V_e = \frac{1}{2}x_1^2 + \frac{1}{2}(x_1 + x_2)^2$$

with

$$\dot{V}_e = x_1 \arctan x_2 + (x_1 + x_2)(u + \arctan x_2)$$

We get

$$\dot{V}_e = -x_1 \arctan x_1 - (x_1 + x_2)^2$$

by taking

$$u = -\arctan x_2 - (x_1 + x_2) - \frac{x_1(\arctan x_1 + \arctan x_2)}{x_1 + x_2}$$

■

Of course the backstepping approach can be used repeatedly. In this way systems with the structure

$$\begin{aligned}
\dot{x} &= f(x, z_1) \\
\dot{z}_1 &= a_1(x, z_1) + b_1(x, z_1) z_2 \\
\dot{z}_2 &= a_2(x, z_1, z_2) + b_2(x, z_1, z_2) z_3 \\
&\vdots \\
\dot{z}_n &= a_n(x, z_1, \dots, z_n) + b_1(x, z_1, \dots, z_n) u
\end{aligned}$$

can be handled, provided a control law and corresponding Lyapunov function is known for the system $\dot{x} = f(x, u)$.

6.3 Forwarding

There is also a method for extension of a Lyapunov function forwards, across a nonlinear integrator. The system structure is assumed to be

$$\dot{z} = a(x) + b(x)u \tag{6.15}$$

$$\dot{x} = f(x) + g(x)u \tag{6.16}$$

There is assumed to be an equilibrium at the origin so that $a(0) = 0$, $f(0) = 0$. We assume that we have a positive definite and radially unbounded function V, together with a control law $k(x)$, such that

$$\dot{V} = V_x(x)(f(x) + g(x)k(x)) \tag{6.17}$$

is negative definite. We also assume that k is such that the asymptotic convergence is actually exponential. We consider the system

$$\dot{z} = a(x) + b(x)k(x) \tag{6.18}$$

$$\dot{x} = f(x) + g(x)k(x) \tag{6.19}$$

Define $h(x) = a(x) + b(x)k(x)$ and let $x = \pi(t, x_o)$ be the solution of (6.19) with initial condition x_o. Then z is given by

$$z(t) = z(0) + \int_0^t h(\pi(s, x(0)))\ ds$$

Since $\pi(t, x(0))$ converges to zero exponentially, the integral converges. We can thus define the variable

$$\zeta = z + \int_0^\infty h(\pi(s, x))\ ds \tag{6.20}$$

Using x and ζ as state variables, (6.18), (6.19) can be written

$$\dot{\zeta} = 0 \tag{6.21}$$

$$\dot{x} = f(x) + g(x)k(x) \tag{6.22}$$

Using the Lyapunov function

$$V_e(x, \zeta) = V(x) + \frac{1}{2}\zeta^2 \tag{6.23}$$

it is clear that $\dot{V}_e = \dot{V} \leq 0$. There is no asymptotic stability, however, since $x = 0$ is an equilibrium regardless of ζ. To get asymptotic stability one can write the control in the form $u = k(x) + \tilde{u}$ so that the system description becomes

$$\dot{\zeta} = \tilde{b}(x)\tilde{u} \tag{6.24}$$

$$\dot{x} = f(x) + g(x)k(x) + g(x)\tilde{u} \tag{6.25}$$

where $\tilde{b} = b + (\partial\zeta/\partial x)\, g$. We can then use (6.23) as a control Lyapunov function for $\tilde{u}$. This gives

$$\dot{V}_e = V_x(f + gk) + (V_x g + \zeta\tilde{b})\tilde{u}$$

We can then use the control

$$\tilde{u} = -(V_x g + \zeta\tilde{b})$$

to stabilize the whole system.

Example 6.3 Consider the system

$$\dot{z} = u$$
$$\dot{x} = -x + x^3 + u$$

The subsystem $\dot{x} = -x + x^3 + u$ can be globally asymptotically stabilized using $u = k(x) = -x^3$ and $V = \frac{1}{2}x^2$. With this feedback the system becomes

$$\dot{z} = -x^3$$
$$\dot{x} = -x$$

The new variable becomes

$$\zeta = z + \int_0^\infty (-x^3 e^{-3t})\, dt = z - \frac{x^3}{3}$$

Using x, ζ as state variables and writing $u = -x^3 + \tilde{u}$ gives

$$\dot{\zeta} = (1 - x^2)\tilde{u}$$
$$\dot{x} = -x + \tilde{u}$$

with the control Lyapunov function

$$V_e = \frac{x^2}{2} + \frac{\zeta^2}{2}$$

Since

$$\dot{V}_e = -x^2 + (x + \zeta - \zeta x^2)\tilde{u}$$

we can take

$$\tilde{u} = -(x + \zeta - \zeta x^2) = -x - (z - \frac{x^3}{3})(1 - x^2)$$

In Figure 6.2 the system is simulated. A comparison is made with the controller which has only linear terms, i.e.

$$\tilde{u} = u = -x - z$$

In this case the linear controller does not stabilize the system for the initial condition $z(0) = 3$, and we see an example of finite escape time. ■

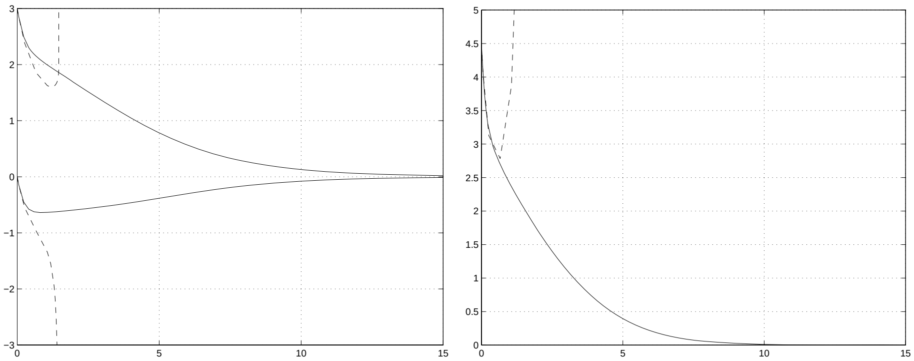


Figure 6.2: System with forwarding controller (solid) compared to a linear controller (dashed). The initial state is $x(0) = 0$, $z(0) = 3$. To the left are the states, to the right the Lyapunov function.

6.4 Stability of observers

Let us now return to the question of stability of the observers discussed in Section 4.3. There we had a system

$$\dot{z} = Az + B\phi(z) + g(z)u, \quad y = Cz \tag{6.26}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & 0 \\ \vdots & & & & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \tag{6.27}$$

and

$$g(x) = \begin{bmatrix} g_1(x_1) \\ g_x(x_1, x_2) \\ \vdots \\ g_{n-1}(x_1, \dots, x_{n-1}) \\ g_n(x) \end{bmatrix} \tag{6.28}$$

Using the natural observer

$$\dot{\hat{x}} = A\hat{x} + B\phi(\hat{x}) + g(\hat{x})u + K(y - C\hat{x}) \tag{6.29}$$

then gives an observer error $\tilde{x} = x - \hat{x}$ that satisfies

$$\dot{\tilde{x}} = (A - KC)\tilde{x} + B(\phi(x) - \phi(\hat{x})) + (g(x) - g(\hat{x}))u = (A - KC)\tilde{x} + L(x, \hat{x}, u) \tag{6.30}$$

where

$$L(x, \hat{x}, u) = B(\phi(x) - \phi(\hat{x})) + (g(x) - g(\hat{x}))\, u$$

Using the Lyapunov function $V = \tilde{x}^T S \tilde{x}$ gives

$$\dot{V} = \tilde{x}^T((A - KC)^T S + S(A - KC))\tilde{x} + 2\tilde{x}^T SL(x, \hat{x}, u)$$

We now choose the gain as $K = S^{-1}C^T$ which gives

$$\dot{V} = \tilde{x}^T(A^T S + SA - 2C^T C)\tilde{x} + 2\tilde{x}^T SL(x, \hat{x}, u)$$

Finally we define S to be the solution of the equation

$$A^T S + SA - C^T C = -\theta S \tag{6.31}$$

where θ is a parameter that is to be chosen sufficiently large. The resulting expression for $\dot{V}$ becomes

$$\dot{V} = -\theta\tilde{x}^T S\tilde{x} - (C\tilde{x})^2 + 2\tilde{x}^T SL(x, \hat{x}, u) \leq -\theta\tilde{x}^T S\tilde{x} + 2\tilde{x}^T SL(x, \hat{x}, u) \tag{6.32}$$

Let $|x|$ denote the ordinary Euclidian norm of x and use the notation $|x|_S = |S^{1/2}x| = (x^T Sx)^{1/2}$. Then (6.32) becomes

$$\frac{d}{dt}(|\tilde{x}|_S^2) \leq -\theta|\tilde{x}|_S^2 + 2|\tilde{x}|_S|L(x, \hat{x}, u)|_S$$

which can be simplified to

$$\frac{d}{dt}(|\tilde{x}|_S) \leq -\frac{\theta}{2}|\tilde{x}|_S + |L(x, \hat{x}, u)|_S \tag{6.33}$$

To go further it is necessary to consider in detail the solution of (6.31). We begin by looking at an example.

Example 6.4 Consider an observer for a two-dimensional system. Then

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad S = \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix}$$

and equation (6.31) becomes

$$\begin{bmatrix}-1 & s_{11}\\ s_{11} & 2s_{12}\end{bmatrix} = -\theta \begin{bmatrix}s_{11} & s_{12}\\ s_{12} & s_{22}\end{bmatrix}$$

with the solution

$$S = \begin{bmatrix}\theta^{-1} & -\theta^{-2}\\ -\theta^{-2} & 2\theta^{-3}\end{bmatrix}, \quad K = S^{-1}C^T = \begin{bmatrix}2\theta\\ \theta^2\end{bmatrix} \tag{6.34}$$

Now consider the term $L(x, \hat{x}, u)$. We have

$$L = \begin{bmatrix}L_1\\ L_2\end{bmatrix} = \begin{bmatrix}(g_1(x_1) - g_1(\hat{x}_1))u\\ \phi(x) - \phi(\hat{x}) + (g_2(x) - g_2(\hat{x}))u\end{bmatrix}$$

Let us make the following assumptions

$$|u(t)| \le u_o, \quad \text{all } t \tag{6.35}$$

$$|(g_1(x_1) - g_1(\hat{x}_1))| \le \tilde{\Lambda}|x_1 - \hat{x}_1| \tag{6.36}$$

$$|(g_2(x) - g_2(\hat{x}))| \le \tilde{\Lambda}_2|x - \hat{x}| \tag{6.37}$$

$$\phi(x) - \phi(\hat{x}) \le \tilde{\Lambda}_3|x - \hat{x}| \tag{6.38}$$

Conditions (6.36) – (6.38) are required to hold for all values of x and $\hat{x}$. This means that we require the functions to satisfy a *global Lipschitz condition*. When this holds together with the bound on u it follows immediately that

$$\begin{aligned} |L_1(x_1, \hat{x}_1, u)| &\le \Lambda_1|\tilde{x}_1| \\ |L_2(x, \hat{x}, u)| &\le \Lambda_2|\tilde{x}| \end{aligned}$$

for some constants Λ_1 and Λ_2. It is now possible to estimate the size of the term $|L|_S$ in (6.33):

$$\begin{aligned} |L|_S = (s_{11}L_1^2 + 2s_{12}L_1L_2 + s_{22}L_2^2)^{1/2} \le \\ (\theta^{-1}\Lambda_1^2|\tilde{x}_1| + 2\theta^{-2}\Lambda_1\Lambda_2|\tilde{x}_1||\tilde{x}| + 2\theta^{-3}\Lambda_2^2|\tilde{x}|^2)^{1/2} \le \\ \underbrace{(c_1^2\Lambda_1^2 + 2c_1c_2\Lambda_1\Lambda_2 + 2c_2^2\Lambda_2^2)^{1/2}}_{c_3}|\tilde{x}|_S \end{aligned}$$

where we have used the inequalities

$$|\tilde{x}_1| \le c_1\theta^{1/2}|\tilde{x}|_S, \quad |\tilde{x}| \le c_2\theta^{3/2}|\tilde{x}|_S$$

that are satisfied for some constants c_1 and c_2. It follows that (6.33) can be rewritten as

$$\frac{d}{dt}(|\tilde{x}|_S) \le -\frac{\theta}{2}|\tilde{x}|_S + c_3|\tilde{x}|_S$$

Since the right hand side is negative if θ is large enough it follows that the observer error goes to zero for arbitrary initial conditions. ■

The calculations of the example are easily generalized to the general case which gives the following theorem

Theorem 6.1 Consider the observer (6.29). Assume that there is a constant u_o so that the control signal satisfies $u(t) \leq u_o$ for all t. Also assume that ϕ and the functions $g_j(x_1, .., x_j)$ satisfy global Lipschitz conditions and that the observer gain is $K = S^{-1}C^T$ with S given by (6.31). Then for each θ which is sufficiently large there exists a constant $C(\theta)$ such that

$$|x(t) - \hat{x}(t)| \leq C(\theta)e^{-\theta t/3}|x(0) - \hat{x}(0)| \tag{6.39}$$

i.e the observer error converges to zero with an arbitrarily fast exponential convergence rate.

Proof. It is easy to see that the structure of (6.31) makes it possible to calculate the elements of S recursively, starting with $s_{11} = \theta^{-1}$. One can also show that this matrix is always positive definite. It is also easy to see that the recursive structure means that an element of S will have the form

$$s_{ij} = \frac{s_{ij}^o}{\theta^{i+j-1}}$$

for some constant s_{ij}^o. This means that the estimate of the size of the term $|L|_S$ that was done in Example 6.4 can be done in a completely analogous way for a general system. One can therefore obtain the estimate

$$\frac{d}{dt}(|\tilde{x}|_S) \leq -\frac{\theta}{2}|\tilde{x}|_S + c_3|\tilde{x}|_S$$

from that example also in the general case. Choosing $\theta \geq 6c_3$ gives the estimate

$$\frac{d}{dt}(|\tilde{x}|_S) \leq -\frac{\theta}{3}|\tilde{x}|_S$$

which directly gives (6.39). ■

This theorem gives a strong global convergence result for a general class of observers. It is however worth noting that it depends on a number of conditions.

- The result is global: it holds for any $x(0)$ and any $\hat{x}(0)$. However it requires the system to be in the special form (6.26). In section 4.2 we discussed how a general system could be brought into this form ba a variable change. For the convergence of the observer to be relevant in the original physical variables this coordinate change has to be global.
- The functions $\phi(x)$ and $g_i(x)$ have to satisfy global Lipschitz conditions which is fairly restrictive.
- The observer gains grow to infinity as θ grows to infinity. This observer design therefore usually requires a measurement of y with very low noise level.

The observer and the closed loop system

In control design an observer is usually used to compute and estimate of the state that is used in a controller based on state feedback. Typically the system dynamics is given by

$$\dot{x} = f(x) + g(x)u \tag{6.40}$$

and we find a control law $u = k(x)$ and a Lyapunov function V such that

$$\dot{V} = V_x(x)(f(x) + g(x)k(x)) \leq -q(x) \leq 0 \tag{6.41}$$

Then an observer is designed that gives a state estimate $\hat{x}$ and the controller

$$u = k(\hat{x}) = k(x - \tilde{x}) \tag{6.42}$$

is used. Suppose a Lyapunov function $V_e(\tilde{x})$ is known for the estimation error (like $|\tilde{x}|_S^2$ in Theorem 6.1) with an inequality

$$\dot{V}_e \leq -q_e(\tilde{x}) \leq 0 \tag{6.43}$$

Then it is natural to try $W = V + V_e$ as a Lyapunov function candidate. One gets

$$\dot{W} \leq -q(x) - q_e(\tilde{x}) + \underbrace{V_x(x)(k(x - \tilde{x}) - k(x))}_{\delta} \tag{6.44}$$

For a specific design it is often possible to show that the term δ is sufficiently small so that the overall expression becomes negative. It is however difficult to give general conditions that guarantee that this will happen.

One might think that the high gain observer described by Theorem 6.1 would guarantee stability also for the closed loop system, since according to (6.39), after an arbitrarily short time the difference between x and $\hat{x}$ becomes negligible. The catch is that there might be very large transients during this short time, due to the large gain in the observer. These large transients could, via the control law destabilize the closed loop system beyond recovery. A high gain observer therefore has to be combined with some scheme for handling these initial transients. Provided this is done it is indeed possible to prove closed loop stability for some classes of systems where nonlinear state feedback is combined with high gain observers.

6.5 Rejection of disturbances

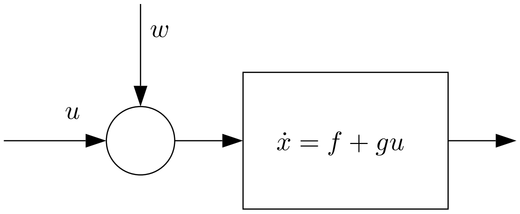
Suppose we have a system

$$\dot{x} = f(t, x) + g(t, x)u \tag{6.45}$$

for which we know a Lyapunov function $V(t, x)$ which guarantees stability for the open loop system by satisfying the conditions of Theorem 5.2. In particular there is a positive definite function W such that

$$V_t + V_x f(t, x) \leq -W(x) \tag{6.46}$$

Now we consider disturbances and uncertainties that can be viewed as additive to the input:



Thus we can write

$$\dot{x} = f(t,x) + g(t,x)(u + w(t,x,u)) \tag{6.47}$$

where we have emphasized that the disturbance is allowed to depend on both the state and the input as well as being time-varying. Disturbances that enter as in (6.47) are sometimes said to satisfy a *matching condition.* If w were known it would of course be possible to eliminate it completely by using feed-forward. We will assume that w is completely unknown except for an upper bound:

$$|w(t,x,u)| \le \rho(t,x) + \gamma|u| \tag{6.48}$$

Computing the derivative of V gives

$$\dot{V} = V_t + V_x(f + gu + gw) \le -W + V_x gu + |V_x g|(\rho + \gamma|u|)$$

Choose u to always give a negative contribution:

$$u = -k(t,x) \text{ sign } (V_x g)$$

It follows that

$$\dot{V} = \le -W + |V_x g|(-k + \rho + \gamma k) \le -W$$

if k is chosen as $\rho/(1-\gamma)$. The chosen control is thus

$$u = -\frac{\rho(t,x)}{1-\gamma} \text{ sign } (V_x g) \tag{6.49}$$

and it achieves that the Lyapunov function decreases at least as fast in the presence of the unknown disturbance as for the open loop undisturbed system. In particular we achieve stability for the closed loop system according to Theorem 5.2. The price we have to pay is that the control law is discontinuous (unless $V_x g = 0$ implies $\rho = 0$) and that there is no guarantee that it will decrease to zero as the equilibrium is approached.

Example 6.5 Consider the mechanical system of figure 6.3. A unit mass with position x_1 and velocity x_2 is connected to a nonlinear spring and a damper. The control signal u is a force and there is also an external force disturbance w. The spring force is assumed to be $-x_1 - x_1^3$ and the damper force is supposed to be $-x_2$. The system equations are then

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - x_1^3 - x_2 + u + w\end{aligned}$$

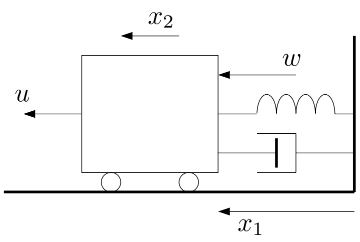


Figure 6.3: Mechanical system with nonlinear spring and force disturbance.

A Lyapunov function for the undisturbed, uncontrolled system, $u = 0$, $w = 0$, is

$$V = \frac{3}{2}x_1^2 + x_1x_2 + x_2^2 + \frac{x_1^4}{2}, \;\Rightarrow\; \dot{V} = -x_1^2 - x_2^2 - x_1^4$$

Suppose we know that $|w| \leq 1$. Then (6.49) gives

$$u = -\operatorname{sign}(V_x g) = -\operatorname{sign}(x_1 + 2x_2)$$

In Figure 6.4 the response of this controller is shown. The disturbance and control signals are shown in Figure 6.5. Note that the control signal switches very rapidly between its extreme values. In Figure 6.6 a low pass filtered version of u is compared to the disturbance. Note that the average of u is close to $-w$, which accounts for the disturbance rejection. ■

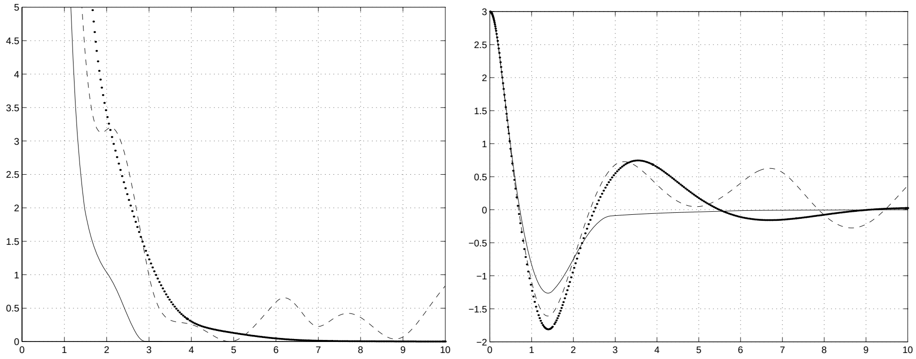


Figure 6.4: Response of the mechanical system. To the left is $V(t)$, to the right $x_1(t)$. Uncontrolled system without disturbance: dotted, uncontrolled system with disturbance: dashed, controlled system with disturbance: solid.

6.6 Passivity based control

In the previous chapter (section 5.5) we discussed passive systems. Passive systems have several interesting properties. One of them is that passivity is preserved under a feedback of the form shown in Figure 6.7.

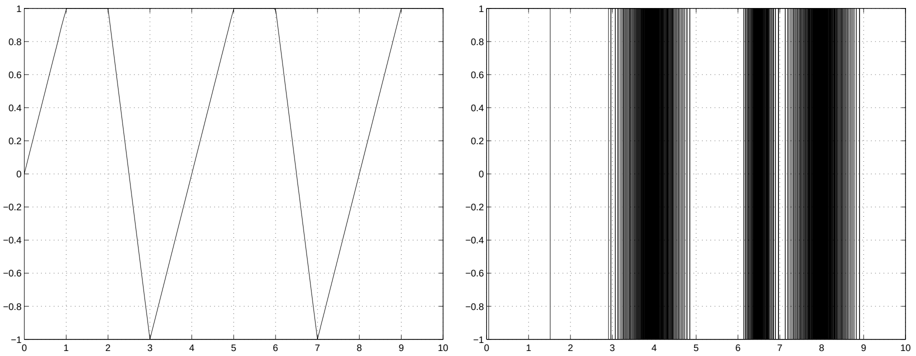


Figure 6.5: Control of mechanical system. Disturbance signal w to the left and control signal u to the right.

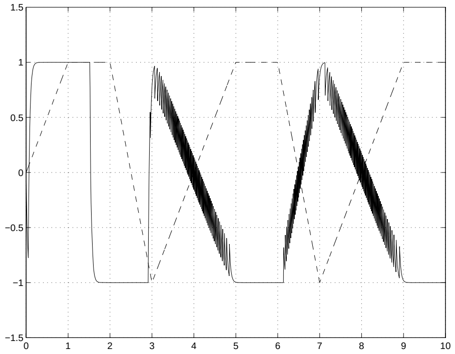


Figure 6.6: Control of mechanical system. Filtered u (solid) and w (dashed).

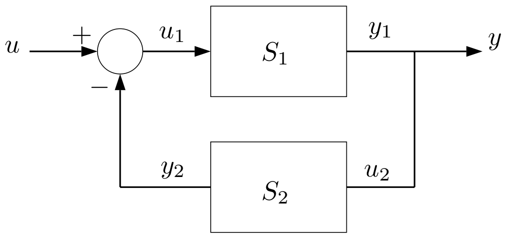


Figure 6.7: Feedback loop with two passive systems.

Proposition 6.1 Consider the feedback system of Figure 6.7 where the systems S_1 and S_2 are both passive. Then the closed-loop system with input u and output y is also passive.

Proof. Since the systems S_1 and S_2 are passive the definition of passivity gives

$$\int_0^T u_i^T y_i \, dt + \gamma_i(x_i^o) \geq 0, \quad i = 1, 2$$

for some funktions γ_1 and γ_2, where x_1^o, x_2^o are the initial states of S_1 and S_2. Using that

$$u_1 = u - y_2, \quad u_2 = y_1 = y$$

one gets

$$\int_0^T u^T y \, dt = \int_0^T (u_1^T y_1 + y_2^T u_2) \, dt \geq -\gamma_1(x_1^o) - \gamma_2(x_2^o)$$

which shows that the system with input u and output y is passive with $\gamma = \gamma_1 + \gamma_2$. ■

In passivity based control one tries to control systems that are already passive in a way that preserves passivity. The proposition above shows that one way of doing this is to use a controller which is in itself passive. Usually this is too restrictive however. Looking at Example 5.5 we see that equation (5.41) can be rewritten as

$$\int_0^T uy \, dt + \gamma(x(0)) - \gamma(x(T)) = \int_0^T by^2 \, dt$$

In passivity-based control one assumes this structure for a (possibly multivariable) system:

$$\int_0^T u^T y \, dt + \gamma(x(0)) - \gamma(x(T)) = d \geq 0 \tag{6.50}$$

where γ represents stored energy and d is a dissipation term. One then considers state feedback of the form

$$u = -k(x) + v \tag{6.51}$$

Substituting into (6.50) one obtains

$$\int_0^T v^T y \, dt - \int_0^T k(x)^T y \, dt + \gamma(x(0)) - \gamma(x(T)) = d$$

If the feedback k can be chosen so that

$$\int_0^T k(x)^T y \, dt = \phi(x(T)) - \phi(x(0))$$

for some function ϕ, then the closed loop system with v as new input will be passive. Its energy storage function will be $\gamma + \phi$. The idea is now to choose ϕ so that the new energy function has a minimum at the desired equilibrium of the system. If the natural dissipation d is not sufficient to drive the system towards the equilibrium fast enough, then further terms are introduced into the controller to increase the dissipation.

This approach has been used extensively for mechanical systems whose stored energy has the form

$$\gamma = \frac{1}{2}\dot{q}^T I(q)\dot{q} + V(q) \tag{6.52}$$

where q is a vector of generalized coordinates (usually linear displacements and angles), I is an inertia matrix and V is the potential energy. Using the feedback

$$k(x) = -V_q + K(q - q_r) \tag{6.53}$$

where K is a positive definite matrix and q_r is the desired position, gives

$$\int_0^T k^T y\, dt = \int_0^T (-V_q + K(q - q_r))\dot{q}\, dt =$$
$$\int_0^T \frac{d}{dt}(-V + \frac{1}{2}(q - q_r)^T K(q - q_r))\, dt = \Phi(x(T)) - \Phi(x(0))$$

where $\Phi = -V + \frac{1}{2}(q - q_r)^T K(g - q_r)$. The stored energy will be modified to

$$\gamma + \Phi = \frac{1}{2}\dot{q}^T I(q)\dot{q} + \frac{1}{2}(q - q_r)^T K(q - q_r)$$

This function has a minimum at $q = q_r$, $\dot{q} = 0$, which is the equilibrium the system will converge to if there is enough dissipation.

Example 6.6 Consider again Example 5.5, but let the damping be a nonlinear function $b(x_2)$ of velocity.

$$\begin{aligned} \dot{x}_1 &= x_2 \\ m\dot{x}_2 &= u - k_1 x_1 - k_2 x_1^3 - b(x_2) \\ y &= x_2 \end{aligned} \tag{6.54}$$

The stored energy is as before

$$\gamma(x) = \frac{k_1}{2}x_1^2 + \frac{k_2}{4}x_1^4 + \frac{m}{2}x_2^2$$

Comparing with (6.52) we see that $q = x_1$, $\dot{q} = x_2$, $I(q) = m$ and $V = \frac{k_1}{2}x_1^2 + \frac{k_2}{4}x_1^4$. The control law (6.51), (6.53) then becomes

$$u = v + k_1 x_1 + k_2 x_1^3 - K(x_1 - r)$$

where r is the reference value for the position. Note that this control law is similar to exact linearization because the spring force is subtracted away. Note also that it is different in not cancelling the damping term $b(x_2)$. The closed loop system will be

$$\begin{aligned} \dot{x}_1 &= x_2 \\ m\dot{x}_2 &= v - K(x_1 - r) - b(x_2) \\ y &= x_2 \end{aligned} \tag{6.55}$$

and its stored energy function

$$V = \frac{m}{2}x_2^2 + \frac{K}{2}(x_1 - r)^2$$

For $v = 0$ and $r =$ constant the closed-loop system has an equilibrium at $x_1 = r$. Using V as a Lyapunov function candidate gives

$$\dot{V} = x_2(-K(x_1 - r) - b(x_2)) + K(x_1 - r)x_2 = -x_2 b(x_2)$$

If the damping satisfies $x_2 b(x_2) \geq 0$, then the equlilibrium $x_1 = r$ is stable. If $x_2 b(x_2) > 0$ for $x_2 \neq 0$ it is asymptotically stable. If the term $-x_2 b(x_2)$ does not give fast enough convergence towards the equilibrium, then an additional, x_2-dependent term in the controller might be useful. ■

6.7 An example from adaptive control

An area where Lyapunov theory has been used extensively is in the design of adaptive controllers. The subject is large and we will just look at a simple example to show the type of reasoning that can be used. Consider a linear system controlled by a P-regulator when the set point is zero and there is an external disturbance w, Figure 6.8. Let the relative degree of G (the difference between the degrees of denominator and numerator) be one, and assume that all zeros are strictly in the left half plane. Then it is easy to see from a root locus argument that G will always be stabilized if k is big enough. Also the influence of a constant w can be made arbitrarily small if k is large enough. However, since G and w are unknown, we do not know how to pick k. A simple idea is to use the adaptation rule

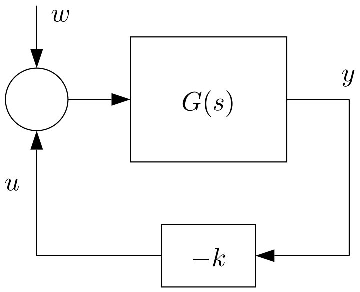


Figure 6.8: Linear system with P-controller and external disturbance.

$$\dot{k} = y^2, \quad k(0) \geq 0$$

which lets k increase as long as there is a control error. This scheme can be shown to work in an ideal noise free situation. In a realistic situation, however, y does not decrease to zero due to disturbances and measurement noise, which means that k will increase without bound. It is clear that some modification is needed, such as

$$\dot{k} = y^2 + f(k), \quad k(0) \geq 0 \tag{6.56}$$

where k has to be determined. We assume that $f(k)$ is chosen so that k always remains positive. To get an idea how the choice should be made, we will try to construct a Lyapunov function. Let n be the order of the linear system G. Since the relative degree is one, there is a state space realization of the form

$$\dot{x} = Ax + by \tag{6.57}$$

$$\dot{y} = -cx - dy + g(u + w) \tag{6.58}$$

with $g \neq 0$, where x is an $(n-1)$-dimensional vector. We assume for simplicity that the problem is scaled so that $g = 1$. Note that it is possible to keep y identically zero by choosing $u = cx + dy$ (for the case $w = 0$). The remaining dynamics is then the zero dynamics, which is given by $\dot{x} = Ax$. It is easy to see that the eigenvalues of A are the zeros of G. Since we assumed G to have its zeros strictly in the left half plane, A has all its eigenvalues there. We will now consider a Lyapunov function candidate of the form

$$V = x^T Px + y^2 + (k - k_o)^2 \tag{6.59}$$

where k_o is a constant to be determined. The analysis will be made for the case of a constant but unknown w. Differentiating we get

$$\dot{V} = (Ax+by)^T Px + x^T P(Ax+by) + 2y(-cx-dy-ky+w) + 2(k-k_o)(y^2+f(k)) =$$

$$= \begin{bmatrix} x^T & y \end{bmatrix} \underbrace{\begin{bmatrix} A^T P + PA & Pb - c^T \\ b^T P + c & -2(d + k_o) \end{bmatrix}}_{Q} \begin{bmatrix} x \\ y \end{bmatrix} + 2wy + 2(k - k_o)f(k)$$

We see that there is no hope of making $\dot{V}$ negative all the time due to the term $2wy$. However, we can do the following. Choose P as the solution of the Lyapunov equation

$$A^T P + PA = -I$$

This is always possible since A has all its eigenvalues strictly in the left half plane. Next choose k_o such that

$$Q = \begin{bmatrix} -I & Pb - c^T \\ b^T P + c & -2(d + k_o) \end{bmatrix}$$

is negative definite. Further choose $f(k)$ such that

$$f(k) \leq -\sigma k, \quad \sigma > 0 \tag{6.60}$$

We then have the following fact.

Proposition 6.2 For the Lyapunov function candidate (6.59) and an adaptation rule satisfying (6.60) there are constants $V_0 > 0$ and $\epsilon > 0$ such that in the set

$$\{(x, y, k) : V(x, y, k) \geq V_0\} \tag{6.61}$$

we have

$$\dot{V} \leq -\epsilon(x^T x + y^2 + k^2)$$

Proof. Since Q is negative definite, there is a constant $\epsilon_1 > 0$ such that $Q \leq -2\epsilon_1 I$. For $k > k_o$ we have

$$\dot{V} \leq -2\epsilon_1(x^T x + y^2) - 2\sigma k^2 + 2\sigma k k_o + 2wy$$

Taking $\epsilon = \min(\epsilon_1, \sigma)$ we get

$$\dot{V} \leq -\epsilon(x^T x + y^2 + k^2)$$

if $(x^T x + y^2 + k^2)$ is large enough, which is the case if $V(x) \geq V_0$ with V_o large enough. ■

The proposition shows that V will decrease in the set $V \geq V_0$ so that eventually the set $V(x, y, k) < V_0$ is reached. The adaptation will thus work in the sense that all variables reach this bounded set. A closer look at the estimates is taken in Exercise 6.10.

6.8 Exercises

6.1 Compute a backstepping controller for the system

$$\begin{aligned} \dot{x}_1 &= x_1^2 + x_2 \\ \dot{x}_2 &= -x_2 + x_3 \\ \dot{x}_3 &= u \end{aligned}$$

6.2 Consider the heat exchanger of example 1.2.

$$\frac{d}{dt}(CT) = qcT_0 - qcT + \kappa(T_h - T)$$

Let the state variables be $x_1 = T$, $x_2 = q$ and $T_h = x_3$. Suppose x_2 and x_3 are controlled from the inputs u_1 and u_2 with some lag due to time constants in the control actuators. If $T_0 = 0$, $c/C = 1$ and $\kappa/C = 1$, then the system is described by

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_3 - x_2 x_1 \\ \dot{x}_2 &= -x_2 + u_1 \\ \dot{x}_3 &= -x_3 + u_2 \end{aligned}$$

Let the control be of the form $u_1 = 1 + \tilde{u}_1$, $u_2 = 1 + \tilde{u}_2$. Compute the equilibrium corresponding to $\tilde{u}_1 = 0$, $\tilde{u}_2 = 0$. Then compute a Lyapunov based feedback that makes that equilibrium asymptotically stable. Is it possible to achieve global asymptotic stability (global in the sense of physically reasonable values of the state variables)?

6.3 The Kokotovic benchmark problem.

$$\begin{aligned} \dot{x}_1 &= x_2 + (x_2 - x_3)^2 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= u \end{aligned}$$

a. Show that the system is not feedback linearizable.
b. Give a control law with associated Lyapunov function that achieves global asymptotic stability.

6.4 Compute a controller for

$$\begin{aligned}\dot{x}_1 &= \sin x_2\\ \dot{x}_2 &= u\end{aligned}$$

with the largest possible stability region (guaranteed by a Lyapunov function).

6.5 Extend the results of the previous problem to

$$\begin{aligned}\dot{x}_1 &= \sin x_2\\ \dot{x}_2 &= \sin x_3\\ \dot{x}_3 &= u\end{aligned}$$

6.6 Solve (6.31) for a third order system and compute the observer gain K. Alos compute the eigenvalues of $A - KC$.

6.7 Consider the system and controller

$$\dot{x} = f(x,u), \quad u = k(x)$$

with the Lyapunov function V satisfying

$$\begin{gathered}c_1|x|^2 \le V(x) \le c_2|x|^2\\ V_x(x)f(x,k(x)) \le -c_3|x|^2, \quad |V_x| \le c_4|x|\end{gathered}$$

for some positive constants c_i. Now suppose there is a disturbance signal w

$$\dot{x} = f(x,k(x)) + w, \quad |w| \le c_5$$

Show that x will eventually satisfy a bound

$$|x| \le c_6$$

Compute an estimate of c_6 based on the other c_i.

6.8 A dissipative system is called strictly input passive if

$$w(u,y) = u^Ty - \epsilon u^Tu, \quad \epsilon > 0$$

and strictly output passive if

$$w(u,y) = u^Ty - \delta y^Ty, \quad \delta > 0$$

Consider the feedback system of Figure 6.7, where S_1 and S_2 are passive. State a condition on strict input or output passivity for the individual systems which guarantees strict output passivity for the closed loop system.

6.9 Simulate the adaptive scheme of Section 6.7 with $f(k) = -\sigma k$ for different systems G and different, not necessarily constant, disturbances w.

6.10 Consider the adaptive scheme of Section 6.7 with $f(k) = -\sigma k$. Let $\sigma < 1/\lambda_{max}(P)$. Show that k_o can be chosen such that

$$\dot{V} \le -2\sigma V + 2\sigma k_o^2 + 2\sigma|w|$$

What conclusions can be drawn about the properties of the adaptive scheme?

.

Chapter 7

Nonlinear optimal control.

In linear system theory many design methods for controllers are based on solutions to the linear quadratic control problem:

$$\min \int_0^{t_1} \left(x^T Q x + u^T R u\right)\, dt + x(t_1)^T Q_0 x(t_1) \tag{7.1}$$

for the system

$$\dot{x} = Ax + Bu \tag{7.2}$$

This chapter deals with the extension of the linear quadratic control ideas to nonlinear systems and non-quadratic criteria.

7.1 The optimal control problem.

We generalize the linear quadratic problem (7.1), (7.2) in the following way. Consider a nonlinear system

$$\frac{d}{dt}x = f(t, x, u), \quad u \in U \tag{7.3}$$

where x is an n-vector and u an m-vector. The control is assumed to be constrained in such a way that $u(t) \in U$ for all t, where U is some set in R^m. Let

$$\pi(t, x_0, u(.))$$

be the solution of (7.3) with the initial condition $x(0) = x_0$ when the control function $u(.)$ is applied. Consider the problem of controlling the system from a given point until a certain condition is met:

$$x(t_0) = x_0 \tag{7.4}$$

$$(t_1, \pi(t_1, x_0, u(.))) \in M \tag{7.5}$$

where M is some set in $R \times R^n$. Typically M might be a point, in which case one would want to reach a certain state at a certain time. As another example,

M might be a line parallel to the time axis, in which case one wants to reach a certain point at an unspecified time. Also we want to find the "best" control satisfying (7.5), so we consider minimizing

$$J = \int_{t_0}^{t_1} L(t, x(t), u(t))dt + \phi(t_1, x(t_1)) \tag{7.6}$$

where we have used the shorthand notation $x(t) = \pi(t, x_0, u(.))$. The problem is thus to transfer the state from the point x_0 to the set M, using controls that lie in U at each instant of time, in such a way that the criterion (7.6) is minimized. Note that the final time t_1 is not specified, except for the conditions imposed by M.

Since we would like the control expressed as a feedback law, we want to solve the problem for all starting times and all starting points. Then J is a function of these quantities as well as the control function:

$$J = J(t_0, x_0, u(.))$$

Suppose a minimizing u exists and define

$$V(t, x) = \min_{u(.)} J(t, x, u(.)) \tag{7.7}$$

(where the minimization is carried out over those functions u that satisfy $u(t) \in U$ and give a trajectory satisfying $(t_1, x(t_1)) \in M$). Thus $V(t, x)$, *the optimal return*, shows the cost of starting from x at time t when the control is chosen optimally.

If we only consider trajectories that reach M once, then from the definition it follows that

$$V(t, x) = \phi(t, x), \quad (t, x) \in M \tag{7.8}$$

The basis for the analysis of the optimal control problem is the following simple relation, sometimes denoted *"the principle of optimality"*.

Theorem 7.1 Let V be defined by (7.7) . Then

$$V(t, x) \leq \int_t^{t+h} L(t, \pi(t, x, u(.)), u(t))dt + V(t+h, \pi(t+h, x, u(.))) \tag{7.9}$$

for all choices of $u(\tau), t \leq \tau \leq t+h$ satisfying $u(\tau) \in U$. Equality is obtained precisely when u is optimal.

Proof. Suppose that an arbitrary $u \in U$ is used from t to $t+h$ and the optimal one from $t+h$ to t_1. Then, letting * denote optimal values

$$\begin{aligned} J(t, x, u(.)) = \int_t^{t+h} & L(t, \pi(t, x, u(.)), u(t))dt + \\ & \int_{t+h}^{t_1} L(t, x^*(t), u^*(t))dt + \phi(t_1, x^*(t_1)) = \\ & \int_t^{t+h} L(t, \pi(t, x, u(.)), u(t))dt + V(t+h, \pi(t+h, x, u(.))) \end{aligned}$$

where $x^*(t) = \pi(t, \pi(t, x, u(.)), u^*(.))$. Since $V(t, x) \leq J(t, x, u(.))$ with equality when u is optimal, the theorem follows. ∎

If the optimal return function is smooth it is possible to go one step further.

Theorem 7.2 If V is continuously differentiable, then it satisfies the following partial differential equation.

$$0 = \min_{u \in U}(V_t(t, x) + V_x(t, x) f(t, x, u) + L(t, x, u)) \tag{7.10}$$

the so called *Hamilton-Jacobi equation* .

Proof. Letting h tend to zero in Theorem 1 gives

$$\frac{d}{dt} V(t, x(t)) + L(t, x(t), u(t)) \geq 0$$

with equality when u is chosen optimally. Using the differentiation rule

$$\frac{d}{dt} V(t, x(t)) = V_t(t, x(t)) + V_x(t, x(t)) f(t, x(t), u(t))$$

then proves the theorem. ∎

The converse of Theorem can also be shown.

Theorem 7.3 Let W be a continuously differentiable function solving the problem

$$\begin{aligned} 0 &= \min\nolimits_{u \in U}(W_t(t, x) + W_x(t, x) f(t, x, u) + L(t, x, u)) \\ W(t, x) &= \phi(t, x), \quad (t, x) \in M \end{aligned} \tag{7.11}$$

Also assume that the minimizing u is given by a function

$$u = k(t, x) \tag{7.12}$$

which is continuous and gives a trajectory satisfying $(t_1, x(t_1)) \in M$.

Then $u = k(t, x)$ is optimal and

$$W(t, x) = V(t, x)$$

Proof. Define x^* as the solution of

$$\frac{d}{dt} x = f(t, x, k(t, x))$$

$$x(t_0) = x_0$$

and let $u^*(t) = k(t, x^*(t))$. Then

$$
\begin{aligned}
J(t_0, x_0, u^*(.)) = \int_{t_0}^{t_1} L(t, x^*(t), u^*(t))dt + \phi(t_1, x^*(t_1)) = \\
\int_{t_0}^{t_1} L(t, x^*(t), u^*(t))dt + W(t_1, x^*(t_1)) = \\
\int_{t_0}^{t_1} \left[L(t, x^*(t), u^*(t)) + \frac{d}{dt} W(t, x^*(t)) \right] dt + W(t_0, x(t_0)) = \\
\int_{t_0}^{t_1} \Big(L(t, x^*(t), u^*(t)) + W_x(t, x^*(t)) f(t, x^*(t), u^*(t)) + W_t(t, x^*(t)) \Big) dt + \\
W(t_0, x(t_0)) = W(t_0, x(t_0)) \leq \\
\int_{t_0}^{t_1} [L(t, x(t), u(t)) + W_t(t, x(t)) + W_x(t, x(t)) f(t, x(t), u(t))] \, dt + \\
W(t_0, x(t_0)) = \int_{t_0}^{t_1} L(t, x(t), u(t))dt + W(t_1, x(t_1)) = \\
\int_{t_0}^{t_1} L(t, x(t), u(t))dt + \phi(t_1, x(t_1)) = J(t_0, x_0, u(.))
\end{aligned}
$$

where u is an arbitrary control signal satisfying $u(t) \in U$ and $(t_1, x(t_1)) \in M$. Consequently u^* is optimal. Since $J(t_0, x_0, u^*(.)) = W(t_0, x(t_0))$ and u^* is optimal, $V = W$. ■

This theorem shows that *if* we can solve (7.11) *and if* the resulting solution W is smooth, then the optimal control problem defined by (7.6) and (7.3) is solved, and the solution is in the feedback form (7.12). Unfortunately the function V defined by (7.7) has a discontinuous gradient in many cases. The theory then becomes much more complicated, since a generalized concept of derivatives is needed to interpret the expression $V_x f$. Even when there is a smooth solution of (7.11), it can seldom be expressed analytically. A simple example where (7.11) can solved explicitly is the following.

Example 7.1 Let the system be

$$\dot{x} = u$$

with the criterion

$$J = \int_{t_0}^{1} \frac{u^4}{4} dt + \frac{x(1)^4}{4}$$

i.e. the final time is specified to be $t_1 = 1$. There is no restriction on the final state or on the control. The Hamilton-Jacobi equation becomes

$$0 = \min_u (V_t + \frac{u^4}{4} + V_x u)$$

For $x \in M$, i.e. for $t_1 = 1$ we have

$$V(1, x) = x^4/4$$

This gives the solution

$$V(t,x) = \frac{x^4}{4(2-t)^3}$$

and the feedback law

$$u = -\frac{x}{2-t}$$

$$\blacksquare$$

7.2 Infinite horizon optimal control.

In the linear quadratic control problem one often considers control over an infinite time horizon.

$$\min \int_0^\infty \left(x^T Q x + u^T R u\right)\, dt \tag{7.13}$$

Analogously it is natural to replace the criterion (7.6) with

$$J = \int_0^\infty L(x(t), u(t)) dt \tag{7.14}$$

Here we have assumed that there is no explicit time dependence in the integrand L. It is then no restriction to assume that the initial time is 0. We also assume that the system is time invariant and described by

$$\frac{d}{dt} x = f(x,u), \quad u \in U \tag{7.15}$$

The optimal return must then also be time invariant: $V(t,x) = V(x)$. The Hamilton-Jacobi equation then reduces to

$$0 = \min_{u \in U} \left(L(x,u) + V_x(x) f(x,u)\right) \tag{7.16}$$

We can now reformulate Theorem 7.7 as

Theorem 7.4 Let W be a continuously differentiable function solving the problem

$$0 \;\; = \;\; \min\nolimits_{u \in U} \left(W_x(x) f(x,u) + L(x,u)\right) \tag{7.17}$$

Also assume that the minimizing u is given by a continuously differentiable function

$$u = k(x) \tag{7.18}$$

that drives the state to the origin as time goes to infinity. Then this control is optimal among all controls that drive the state to the origin, and W is the corresponding optimal return function.

Proof. Consider a control u driving the corresponding x to the origin as t goes to infinity. Then

$$J(x_0, u(.)) = \int_0^\infty L(x(t), u(t))dt =$$

$$\int_0^\infty (L(x(t), u(t)) + W_x(x(t))f(x(t), u(t)))\, dt - \int_0^\infty \frac{d}{dt} W(x(t)(t))dt =$$

$$= \int_0^\infty (L(x(t), u(t)) + W_x(x(t))f(x(t), u(t)))\, dt + W(x_0) \geq W(x_0)$$

where the last inequality follows from (7.17). It follows that $u = k(x)$ is minimizing and that W is the optimal return. ■

Example 7.2 Consider the system

$$\dot{x} = u, \quad |u| \leq 1$$

with the criterion

$$\int_0^\infty (x^2 + u^2)dt$$

The Hamilton-Jacobi equation becomes

$$0 = \min_{|u|\leq 1} (x^2 + u^2 + V_x u)$$

which has the solution

$$V(x) = \begin{cases} x^2, & |x| \leq 1 \\ x^3/3 + x - 1/3, & |x| > 1 \end{cases}$$

with the feedback

$$u = \begin{cases} -x, & |x| \leq 1 \\ -sgn(x), & |x| > 1 \end{cases}$$

■

7.3 Calculation of optimal control.

Example 7.2 is atypical because it allows an explicit closed form solution. This is not possible to achieve in the general case. We will look att some methods of calculating the optimal control numerically.

Series expansion

If the functions L and f in (7.16) are real analytic (i.e. given by convergent series expansions), then it is possible to compute the solution as a series expansion. To show this we assume that the optimal control problem is given by

$$\min \int_0^\infty \left(l(x) + \frac{1}{2} u^T R u \right) dt, \quad \frac{d}{dt} x = a(x) + b(x)u \tag{7.19}$$

where the functions a, b and l are real analytic. We thus assume that the control appears linearly in the dynamics and quadratically in the criterion. This is not necessary for the calculations that follow, but make them very much simpler. The Hamilton-Jacobi equation takes the form

$$0 = l(x) + V_x(x)a(x) - \frac{1}{2}V_x(x)b(x)R^{-1}b(x)^T V_x(x)^T \tag{7.20}$$

and the optimal control is given by

$$u = k(x) = -R^{-1}b(x)^T V_x(x)^T \tag{7.21}$$

Writing

$$\begin{aligned} l(x) &= \frac{1}{2}x^T Qx + l_h(x) \\ a(x) &= Ax + a_h(x) \\ b(x) &= B + b_h(x) \\ V(x) &= \frac{1}{2}x^T Sx + V_h(x) \end{aligned} \tag{7.22}$$

where l_h,a_h,b_h and V_h contain higher order terms (beginning with degrees 3,2,1 and 3 respectively), the Hamilton-Jacobi equation splits into two equations.

$$0 =Q + A^T S + SA - SBR^{-1}B^T S \tag{7.23}$$

$$0 =V_{hx}(x)A_c x + l_h(x) + V_x(x)a_h(x) - \frac{1}{2}V_{hx}(x)BR^{-1}B^T V_{hx}(x)^T - \tag{7.24}$$

$$-\frac{1}{2}V_x(x)w_h(x)V_x(x)^T \tag{7.25}$$

where

$$A_c = A - BR^{-1}B^T S \tag{7.26}$$

$$w_h(x) = b(x)R^{-1}b(x)^T - BR^{-1}B^T \tag{7.27}$$

(so that w_h contains terms of degree 1 and higher). Equation (7.23) is the ordinary Riccati equation of linear quadratic control. If we assume that

$$Q \geq 0, \;\; R > 0, \;\; (Q, A) \text{ observable}, \;\; (R, A)\text{controllable} \tag{7.28}$$

then the theory of linear quadratic control tells us that (7.23) has a unique positive definite solution. Letting superscript (m) denote m:th order terms, equation (7.25) can be written

$$\begin{gathered} -(V^{(m)})_x A_c x = l^{(m)}(x) + \\ + \left(V_x(x)a_h(x) - \tfrac{1}{2}V_{hx}(x)BR^{-1}B^T V_{hx}(x)^T - \tfrac{1}{2}V_x(x)w_h(x)V_x(x)^T\right)^{(m)} \end{gathered} \tag{7.29}$$

The right hand side contains only $(m-1)$:th, $(m-2)$:th,.. order terms of V. Equation (7.29) therefore defines a linear system of equations for the m:th order coefficients with a right hand side which is known if lower order terms have been computed. Using the same arguments that was used in the calculation of Lyapunov functions, it can be shown that this system of equations is nonsingular as soon as A_c is a stable matrix. Since A_c represents the linearized closed loop dynamics, this is guaranteed by linear quadratic control theory if (7.28) holds. After solving (7.23), the 3rd, 4th, 5th,.. order coefficients can then be computed successively. It can be shown that the resulting series representation of V converges and defines a real analytic function in some neighborhood of the origin.

Model predictive control

Consider again the infinite horizon optimal control problem (7.14), (7.15). If the optimal return V is known it is easy to see that (7.14) is equivalent to the finite horizon optimal control problem

$$J = \int_0^T L(x(t), u(t))dt + V(x(T)) \tag{7.30}$$

This formula is only of theoretical interest, since V is usually unknown. There might however be some function ϕ, which approximates V. One then has the approximating finite horizon problem

$$\tilde{J} = \int_0^T L(x(t), u(t))dt + \phi(x(T)) \tag{7.31}$$

Now suppose that we do the minimization over a restricted class of control functions u, namely those that are piecewise constant:

$$u(t) = u_k, \quad kh \leq t < (k+1)h, \quad k = 0, \ldots N-1 \tag{7.32}$$

Here we assumed that T is a multiple of h, $T = Nh$. Now $\tilde{J}$ becomes a function of N real numbers, u_0, ..., u_{N-1}. To minimize $\tilde{J}$ is thus a problem in nonlinear programming that can be solved using commercial software. The problem is difficult for several reasons. The computation of $\tilde{J}$ for given values of u_i involves solving $\dot{x} = f(x, u)$ numerically and then calculating the integral in (7.31) numerically. The problem is a constrained one, since the constraints on u in (7.15) translates into constraints on the u_i. Since the problem in the general case in non-convex, there is no guarantee that a global minimum is found by a numerical algorithm. However, in situations where the nonlinear programming problem can be solved it forms the basis for *nonlinear MPC (model predictive control)*. In nonlinear MPC one solves the optimization problem on line with the current value of x as the initial state to give u_0,..,u_{N-1}. Then u_0 is used during a time interval of length h. After that the problem is resolved, regarding the current time as $t = 0$ and using the actual measured x as a new initial value of the optimization problem. Again only u_0 is used and then the problem is resolved. Since the current, measured, state x is used as the initial state in the minimization of $\tilde{J}$, this gives a feedback controller. Since at each time (7.31) is solved for a time interval which is T units forward from the present time, this approach is also referred to as a *receding horizon* optimal control. It is intuitively clear that nonlinear MPC will give a good approximation of the solution to the original optimal control problem (7.30) if T is large enough, h small enough and ϕ a good enough approximation of V.

The MPC technique has been particularly successful for the constrained linear quadratic problem

$$\begin{aligned}
&J = \int_0^\infty (x^T Q x + u^T R u)\, dt \\
&\dot{x} = Ax + Bu, \quad u_i^b \le u_i \le u_i^t, \;\; i = 1, \ldots, m \\
&\;\; Nx + Mu \le 0
\end{aligned}$$

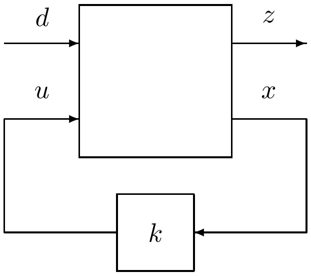
The receding horizon problem then has a criterion of the form

$$\tilde{J} = \int_0^T (x^T Q x + u^T R u)\, dt + x(T)^T Q_o x(T)$$

The finite dimensional optimization problem for u_0,..,u_{N-1} is still quadratic with linear constraints. It can be solved using efficient quadratic programming algorithms.

7.4 Nonlinear robustness techniques

As in the linear case, many robustness problems can be formulated in the following framework:



The external signal d could be a disturbance or a reference signal. The output z is some measure of the control error and should be small. The feedback k then has to be chosen so that the gain γ from d to z is as small as possible. Mathematically we formulate the problem as follows. Consider the system

$$\begin{aligned}
\dot{x} &= f(x) + g(x)u + b(x)d \\
z &= h(x)
\end{aligned} \tag{7.33}$$

where x is an n-vector, u an m-vector, y a p-vector and d a q-vector. We will look at state feedback

$$u = k(x) \tag{7.34}$$

so the resulting closed loop system becomes

$$\begin{aligned}
\dot{x} &= f(x) + g(x)k(x) + b(x)d \\
y &= \begin{bmatrix} z \\ u \end{bmatrix} = \begin{bmatrix} h(x) \\ k(x) \end{bmatrix}
\end{aligned} \tag{7.35}$$

We will look at the following problems

Definition 7.1 The state feedback gain reduction problem is to find a state feedback (7.34) such that the gain from d to y of the closed loop system (7.35) is less than some given value γ.

Definition 7.2 The *optimal* state feedback gain reduction problem is to find the smallest value γ^*, such that the state feedback gain reduction problem is solvable for all $\gamma > \gamma^*$.

To compute the gain from d to y we can use the Hamilton-Jacobi inequality (5.52). We get

$$V_x(f+gk) + \frac{1}{4\gamma^2} V_x bb^T V_x^T + h^T h + k^T k \le 0 \tag{7.36}$$

Completing squares we can also write

$$V_x f + \frac{1}{4\gamma^2} V_x bb^T V_x^T + h^T h + (k + \frac{1}{2} g^T V_x^T)^T (k + \frac{1}{2} g^T V_x^T) - \frac{1}{4} V_x gg^T V_x^T \le 0 \tag{7.37}$$

We get the following result.

Theorem 7.5 Let γ be given. If the Hamilton-Jacobi inequality

$$V_x f + \frac{1}{4} V_x \left(\frac{1}{\gamma^2} bb^T - gg^T \right) V_x^T + h^T h \le 0, \quad V(x_0) = 0 \tag{7.38}$$

has a nonnegative solution V, then the control law

$$u = k(x), \quad k(x) = -\frac{1}{2} g(x)^T V_x(x)^T \tag{7.39}$$

gives a gain from d to y which is less than or equal to γ, that is

$$\int_0^T (z^T z + u^T u)\, dt \le \gamma^2 \int_0^T d^T d\, dt \tag{7.40}$$

for all solutions of the closed loop system starting at x_0.

Solving the Hamilton-Jacobi inequality leads to the same type of problems as the solution of the Hamilton-Jacobi equation. We give a simple example where there is an explicit solution.

Example 7.3 (Van der Schaft's system). Consider the system

$$\dot{x} = u + (\arctan x) d \tag{7.41}$$

$$y = \begin{pmatrix} x \\ u \end{pmatrix} \tag{7.42}$$

The Hamilton-Jacobi inequality becomes

$$\frac{1}{4} \left(\left(\frac{\arctan x}{\gamma} \right)^2 - 1 \right) V_x^2 + x^2 \le 0$$

A global solution requires that

$$|\arctan x| < \gamma$$

that is $\gamma > \pi/2$. A feedback giving gain γ is then

$$u = -\frac{x}{\sqrt{1 - (\arctan x/\gamma)^2}}$$

■

7.5 Exercises.

7.1 What is the optimal feedback for the system

$$\dot{x} = u$$

with the criterion

$$\int_0^\infty \frac{1}{2}(x^2 + x^4 + u^2)dt$$

What is the optimal cost?

7.2 What are the optimal feedback and the optimal return for the system

$$\dot{x} = u$$

with the cost

$$\int_0^\infty \frac{1}{2}(x^4 + u^2)dt$$

Is the optimal return function real analytic?

7.3 Compute the optimal feedback up to third order terms for the system and criterion which are given by

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u$$

$$J = \int_0^\infty \left(x_1^2 + x_1^4 + u^2\right) dt$$

Hint: Maple, Mathematica.

7.4 Consider a constrained linear quadratic control problem with a scalar u:

$$J = \int_0^\infty (x^T Q x + u^2)\, dt$$

$$\dot{x} = Ax + Bu, \quad |u| \leq 1$$

Let $u = -Lx$ be the solution to linear quadratic problem without the constraint $|u| \leq 1$. Then it is natural to try the control law

$$u = -\text{sat}\, Lx$$

where sat is the function given by

$$\operatorname{sat} z = \begin{cases} 1 & z > 1 \\ z & |z| \le 1 \\ -1 & z < -1 \end{cases}$$

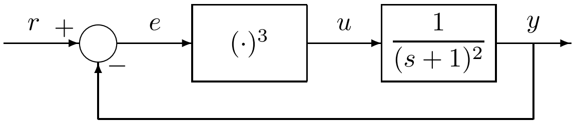
a. Show that this control law is indeed optimal if x is a scalar.
b. Show that the control law is non-optimal in general.

Chapter 8

Harmonic analysis of nonlinear systems

Linear systems are often analyzed by considering their response to sinusoidal signals. For nonlinear systems such an analysis is more complicated since there is no superposition principle. However it is still possible to get some feeling for the system properties from such an analysis. We begin by considering a simple example.

Example 8.1 Consider the following nonlinear control system.



where the nonlinearity $u = e^3$ is followed by the linear system

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

In figure 8.1 is shown the step response for a unit step

$$r(t) = 1, \quad t \geq 0$$

in the reference signal. We see that there is a large steady state error due to the low gain of the cubic nonlinearity at low amplitudes. If the reference signal is instead

$$r(t) = 1 + 1.5 \sin 10t, \quad t \geq 0 \tag{8.1}$$

the response is the one shown in figure 8.2. We see that the steady state error has decreased dramatically due to the presence of a high frequency component in r. Can we explain this phenomenon?

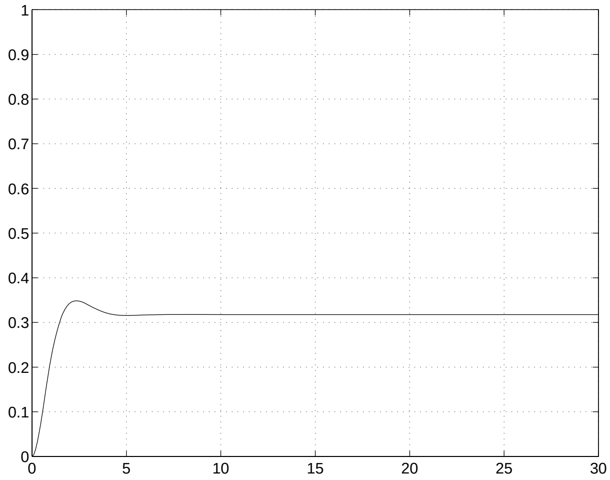


Figure 8.1: Output $y(t)$ for a unit step in the reference.

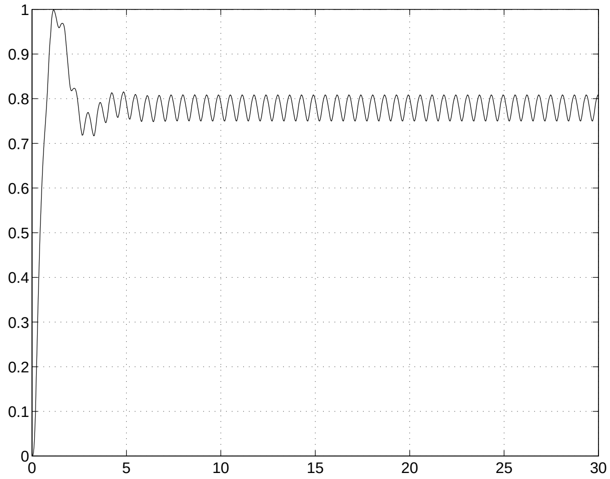


Figure 8.2: Output $y(t)$ for step plus sinusoid.

One approach would be to assume that e is a sum of a constant and a sinusoid. Using a complex representation we write

$$e(t) = e_0 + e_1 e^{i\omega t} + e_{-1} e^{-i\omega t} \tag{8.2}$$

Since $e(t)$ is real we have $e_{-1} = \bar{e}_1$. We then get

$$u(t) = \left(e_0 + e_1 e^{i\omega t} + \bar{e}_1 e^{-i\omega t}\right)^3 = u_0 + u_1 e^{i\omega t} + \bar{u}_1 e^{-i\omega t} + u_2 e^{i2\omega t} + \bar{u}_2 e^{-i2\omega t} + \\ + u_3 e^{i3\omega t} + \bar{u}_3 e^{-i3\omega t}$$

where

$$u_0 = e_0^3 + 6e_0|e_1|^2, \quad u_1 = 3e_0^2 e_1 + 3e_1|e_1|^2, \quad u_2 = 3e_0 e_1^2, \quad u_3 = e_1^3$$

The output is then given by

$$y(t) = G(0)u_0 + G(i\omega)u_1 e^{i\omega t} + G(-i\omega)\bar{u}_1 e^{-i\omega t} + G(2i\omega)u_2 e^{i2\omega t} + \\ + G(-2i\omega)\bar{u}_2 e^{-i2\omega t} + G(3i\omega)u_3 e^{i3\omega t} + G(-3i\omega)\bar{u}_3 e^{-i3\omega t}$$

Since $e = r - y$, the signal e will contain terms with frequencies 2ω and 3ω. Our assumption (8.2) is then false. However, if ω is large enough, then the absolute values of $G(2i\omega)$ and $G(3i\omega)$ will be small. The assumption (8.2) will then be approximately true. Writing the reference signal in the form

$$r(t) = r_0 + r_1 e^{i\omega t} + \bar{r}_1 e^{-i\omega t}$$

we get

$$e_0 = r_0 - G(0)u_0, \quad e_1 = r_1 - G(i\omega)u_1$$

It is convenient to introduce the gains

$$Y_0(e_0, e_1) = \frac{u_0}{e_0} = e_0^2 + 6|e_1|^2, \quad Y_1(e_0, e_1) = \frac{u_1}{e_1} = 3e_0^2 + 3e_1^2 \tag{8.3}$$

We then get

$$e_0 = \frac{r_0}{1 + Y_0(e_0, e_1)G(0)}, \quad e_1 = \frac{r_1}{1 + Y_1(e_0, e_1)G(i\omega)} \tag{8.4}$$

These formulas look superficially like the usual formulas for the gain from reference signal to error signal for a linear system. The big difference lies in the fact that Y_0 and Y_1 are amplitude dependent gains. Note that each gain depends on both e_0 and e_1. It is this fact which is the clue to the behavior of the system.

Substituting (8.3) into (8.4) gives the following system of equations.

$$\begin{aligned} e_0(1 + G(0)(e_0^2 + 6|e_1|^2)) &= r_0 \\ e_1(1 + 3G(i\omega)(e_0^2 + e_1^2)) &= r_1 \end{aligned} \tag{8.5}$$

Since $|G(i\omega)| = |G(10i)| \approx 0.01$ we get approximately

$$e_1 \approx r_1$$

and substituting into the first equation

$$e_0 \approx \frac{r_0}{1 + 6G(0)|r_1|^2} = \frac{r_0}{1 + 6|r_1|^2}$$

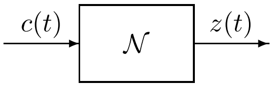
Since we had $r_0 = 1$ and $r_1 = 0.75$ in (8.1), we get

$$e_0 = 0.23$$

in good agreement with Figure (8.2). ■

8.1 Describing functions

Generalizing the methodology of Example 8.1 we consider the following situation.



A nonlinear system $\mathcal{N}$ has an input $c(t)$ and an output $z(t)$. Assume that the input is a sum of sinusoidal functions:

$$c(t) = C_1 \sin(\omega_1 t + \phi_1) + \cdots + C_N \sin(\omega_N t + \phi_N)$$

Using a complex representation we can write

$$c(t) = \sum_{j=1}^{N} \left(c_j e^{i\omega_j t} + \bar{c}_j e^{-i\omega_j t} \right) \tag{8.6}$$

If the output is written in the form

$$z(t) = \sum_{j=1}^{N} \left(z_j e^{i\omega_j t} + \bar{z}_j e^{-i\omega_j t} \right) + w(t) \tag{8.7}$$

where $w(t)$ contains frequencies other than $\omega_1, \ldots, \omega_N$ we can define the gain for the frequency ω_j:

$$Y_j(c_1, \ldots, c_N, \omega_1, \ldots, \omega_N) = \frac{z_j}{c_j} \tag{8.8}$$

The function Y_j is sometimes called the j:th *describing function* for $\mathcal{N}$. Using the gain Y_j it is in principle possible to do block diagram calculations in the same manner that one does for linear systems. We saw this in Example 8.1. There are two main difficulties with this approach.

- In contrast to the linear case, w will in general be nonzero in steady state, that is new frequencies are generated. If $\mathcal{N}$ is part of a feedback loop, these new frequencies will come back at the input. If the new frequency components are included in c, still new frequencies will be generated and so on. Exact calculations will thus require that c contains infinitely many frequencies. To get around this problem one usually assumes that there are linear elements that dampen the frequencies one does not want to consider. In Example 8.1 we did this for the frequencies 2ω and 3ω.

- In general each Y_j depends on all the frequencies ω_j and all the amplitudes c_j. If $\mathcal{N}$ is a static nonlinearity the dependence on the frequencies will dissappear as in Example 8.1. However the dependence on all amplitudes remains. This means that signal components at different frequencies are coupled, which complicates the calculations.

In practice describing function calculations have been confined to the following special cases. They cover many applications however.

1. $\mathcal{N}$ is in fact linear with transfer function $G(s)$. Then trivially
$$Y_j = G(i\omega_j)$$
2. The signal $c(t)$ has zero mean and contains a single frequency, that is
$$c(t) = C \sin \omega t = \frac{C}{2i} e^{i\omega t} - \frac{C}{2i} e^{-i\omega t}$$
If the nonlinearity is static, the single describing function Y depends only on C. This is the classical approach and sometimes the term "describing function method" is used for this special case only.
3. The input signal $c(t)$ has nonzero mean and contains a single frequency (apart from $\omega = 0$), that is
$$c(t) = B + C \sin \omega t = B + \frac{C}{2i} e^{i\omega t} - \frac{C}{2i} e^{-i\omega t}$$
This case is sometimes referred to as the BSDF (bias-plus-sinusoid describing function) case.
4. The signal $c(t)$ has zero mean and contains two frequencies ω_1 and ω_2, that is
$$c(t) = C_1 \sin \omega_1 t + C_2 \sin(\omega_2 t + \phi)$$
This is sometimes referred to as the SSDF case (sinusoid-plus-sinusoid describing function)

8.1.1 The classical describing function

As discussed above, the classical approach assumes a single sinusoid without bias. Usually the nonlinearity is assumed to be static so that

$$z(t) = f(c(t)) \tag{8.9}$$

The classical approach also assumes an odd nonlinearity, that is

$$f(x) = -f(-x) \tag{8.10}$$

The reason is that nonlinearities without this property usually produce a bias at the output. If the nonlinearity is part of a feedback system, this bias is fed back to the input, making the assumption of zero bias invalid. If the static nonlinearity (8.9) is used with the input

$$c(t) = C \sin \omega t = \frac{C}{2i} e^{i\omega t} - \frac{C}{2i} e^{-i\omega t} \tag{8.11}$$

the output becomes a periodic function with a Fourier expansion

$$z(t) = \sum_{-\infty}^{\infty} z_k e^{ik\omega t} = \sum_{-\infty}^{\infty} z_k e^{ik\theta}$$

where $\theta = \omega t$. The Fourier coefficient of the bias term is

$$z_0 = \int_{-\pi}^{\pi} f(C \sin\theta) d\theta = 0$$

because of (8.10). The coefficient of the ω-component is

$$z_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(C \sin\theta) e^{-i\theta} d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(C \sin\theta) \cos\theta d\theta - i \frac{1}{2\pi} \int_{-\pi}^{\pi} f(C \sin\theta) \sin\theta d\theta$$

Since the cosine integral is zero because of the symmetry (8.10), the describing function becomes

$$Y(C) = \frac{z_1}{(\frac{C}{2i})} = \frac{1}{\pi C} \int_{-\pi}^{\pi} f(C \sin\theta) \sin\theta d\theta$$

which is real valued. Since each quarter of the interval $[-\pi, \pi]$ makes the same contribution in the integral, we can also write

$$Y(C) = \frac{4}{\pi C} \int_{0}^{\pi/2} f(C \sin\theta) \sin\theta d\theta \tag{8.12}$$

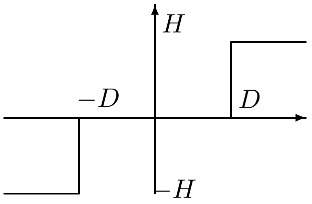
This formula can be rewritten in many ways. Making the substitution $C \sin\theta = x$, for instance gives

$$Y(C) = \frac{4}{\pi C^2} \int_{0}^{C} \frac{x f(x)}{\sqrt{C^2 - x^2}} dx \tag{8.13}$$

Using a partial integration, this can then be written

$$Y(C) = \frac{4}{\pi C^2} \int_{0}^{C} f'(x) \sqrt{C^2 - x^2}\, dx \tag{8.14}$$

Example 8.2 The describing function for a relay with dead zone



is

$$Y(C) = \frac{4}{\pi C^2} \int_{0}^{C} H \;\; \delta(x - D) \;\; \sqrt{C^2 - x^2}\, dx = \frac{4H}{\pi C^2} \sqrt{C^2 - D^2}$$

■

Nonlinearities with hysteresis

A nice feature of the classical describing function method is that it is one of the few methods that can handle nonlinearities that are not single valued. Consider

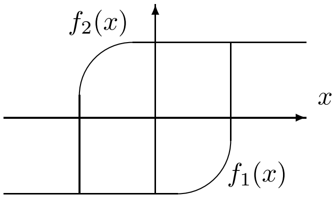


Figure 8.3: Nonlinearity with hysteresis

for instance a nonlinearity with hysteresis as in Figure 8.3. We still consider nonlinearities with an odd symmetry:

$$f_1(x) = -f_2(-x) \tag{8.15}$$

The describing function is then given by

$$Y(C) = \frac{i}{\pi C}\left\{\int_{-\pi/2}^{\pi/2} f_1(C\sin\theta)e^{-i\theta}\,d\theta + \int_{\pi/2}^{-\pi/2} f_2(C\sin\theta)e^{-i\theta}\,d\theta\right\} \tag{8.16}$$

Introducing the mean value

$$f_0(x) = \frac{f_1(x)+f_2(x)}{2}$$

and half the difference

$$\epsilon(x) = \frac{f_2(x)-f_1(x)}{2} = f_2(x) - f_0(x) = f_0(x) - f_1(x)$$

we see that f_0 and ϵ are odd and even functions respectively:

$$f_0(x) = -f_0(-x), \quad \epsilon(x) = \epsilon(-x) \tag{8.17}$$

Replacing f_1 and f_2 in (8.16) by f_0 and ϵ, it is easy to see that the describing function is given by

$$Y(C) = \frac{4}{\pi C}\left\{\int_0^{\pi/2} f_0(C\sin\theta)\sin\theta\,d\theta - i\int_0^{\pi/2} \epsilon(C\sin\theta)\cos\theta\,d\theta\right\} \tag{8.18}$$

Making again the substitution $x = C\sin\theta$ we get the alternative expression

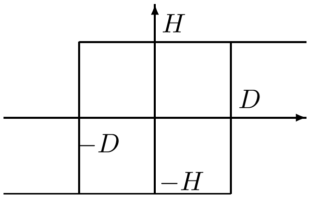
$$Y(C) = \frac{4}{\pi C^2}\left\{\int_0^C f_0'(x)\sqrt{C^2-x^2}\,dx - i\int_0^C \epsilon(x)\,dx\right\} \tag{8.19}$$

Note that the imaginary part of Y is just

$$\frac{A}{\pi C^2}$$

where A is the area enclosed by the hysteresis loop.

Example 8.3 Consider a relay with hysteresis:



The function f_0 equals the nonlinearity of Example 8.2 and the area enclosed by the hysteresis loop is $4HD$. The describing function is thus

$$Y(C) = \frac{4H}{\pi C^2}\sqrt{C^2 - D^2} - i\frac{4HD}{\pi C^2}$$

■

8.1.2 Describing functions for bias plus sinusoid

Now consider an input signal of the form

$$c(t) = B + C\sin\omega t = B + C\sin\theta$$

and a system $\mathcal{N}$ described by a nonlinear function f (not necessarily possessing any symmetries). The output signal is then

$$z(t) = f(B + C\sin\theta) = z_0 + z_1 e^{i\theta} + \bar{z}_1 e^{-i\theta} + \cdots$$

The describing function has two components

$$Y_0(B, C) = \frac{z_0}{B} = \frac{1}{2\pi B}\int_{-\pi}^{\pi} f(B + C\sin\theta)\, d\theta \tag{8.20}$$

and

$$Y_1(B, C) = \frac{z_1}{\frac{C}{2i}} = \frac{1}{\pi C}\int_{-\pi}^{\pi} f(B + C\sin\theta)(\sin\theta + i\cos\theta)\, d\theta \tag{8.21}$$

If we want to cover systems with hysteresis as in Figure 8.3 we can still define

$$f_0(x) = \frac{f_1(x) + f_2(x)}{2}$$

$$\epsilon(x) = \frac{f_2(x) - f_1(x)}{2} = f_2(x) - f_0(x) = f_0(x) - f_1(x)$$

However, since we do not assume any symmetry of f_1 and f_2, we will no longer obtain the symmetries (8.17). Partitioning the integration intervals of (8.20) and (8.21) in a suitable manner we get

$$Y_0(B, C) = \frac{1}{\pi B}\int_{-\pi/2}^{\pi/2} f_0(B + C\sin\theta)\, d\theta \tag{8.22}$$

$$Y_1(B, C) = \frac{2}{\pi C}\left(\int_{-\pi/2}^{\pi/2} f_0(B + C\sin\theta)\sin\theta\, d\theta - i\int_{-\pi/2}^{\pi/2} \epsilon(B + C\sin\theta)\cos\theta\, d\theta\right) \tag{8.23}$$

Making the same variable changes as in (8.12,8.13) we get the alternative formulas

$$Y_0(B,C) = \frac{1}{\pi B}\int_{-C}^{C} \frac{f_0(B+x)}{\sqrt{C^2-x^2}}\,dx \tag{8.24}$$

$$Y_1(B,C) = \frac{2}{\pi C^2}\left(\int_{-C}^{C} \frac{x f_0(B+x)}{\sqrt{C^2-x^2}}\,dx - i\int_{-C}^{C}\epsilon(B+x)\,dx\right) \tag{8.25}$$

or

$$Y_1(B,C) = \frac{2}{\pi C^2}\left(\int_{-C}^{C} f_0'(B+x)\sqrt{C^2-x^2}\,dx - i\int_{-C}^{C}\epsilon(B+x)\,dx\right) \tag{8.26}$$

Example 8.4 Consider the relay with hysteresis of Example 8.3. Let the input have a bias, that is

$$c(t) = B + C\sin\omega t$$

with $C - |B| > D$ (this ensures that the whole hysteresis loop is covered). We get from (8.24)

$$Y_0(B,C) = \frac{1}{\pi B}\left(\int_{-C}^{-B-D} \frac{-H}{\sqrt{C^2-D^2}}\,dx + \int_{-B+D}^{C}\frac{H}{\sqrt{C^2-D^2}}\,dx\right) =$$

$$= \frac{H}{\pi B}\left(\arcsin\left(\frac{D+B}{C}\right) - \arcsin\left(\frac{D-B}{C}\right)\right)$$

Since

$$f_0'(x) = H(\delta(x+D)+\delta(x-D))$$

we get from (8.26)

$$\operatorname{Re} Y_1(B,C) = \frac{2H}{\pi C}\left(\sqrt{1-\left(\frac{D+B}{C}\right)^2} + \sqrt{1-\left(\frac{D-B}{C}\right)^2}\right)$$

$$\operatorname{Im} Y_1(B,C) = \frac{-2}{\pi C^2}\int_{-B-D}^{-B+D} H\,dx = -\frac{4HD}{\pi C^2}$$

■

8.2 Analysis of systems using the describing function

In Example 8.1 we already saw the basic idea of describing function analysis. The various signals are traced around the loop, resulting in a system of equations. We give another example.

Example 8.5 Consider the temperature control system of Figure 8.4. The control signal is the heating power u which can either be switched on (with power $2P$) or off (zero power) by a relay, depending on the difference between the setpoint temperature r and the actual temperature y. The heated object

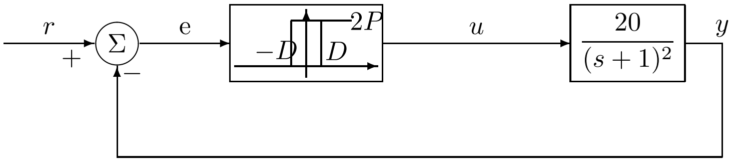


Figure 8.4: Temperature control system

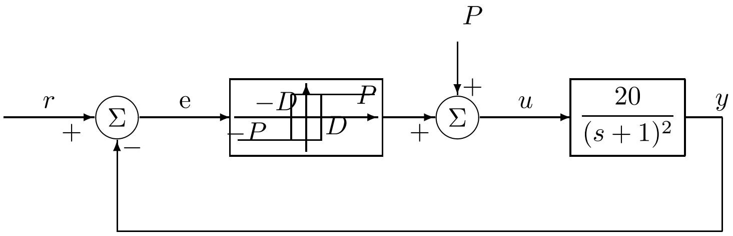


Figure 8.5: Modified block diagram

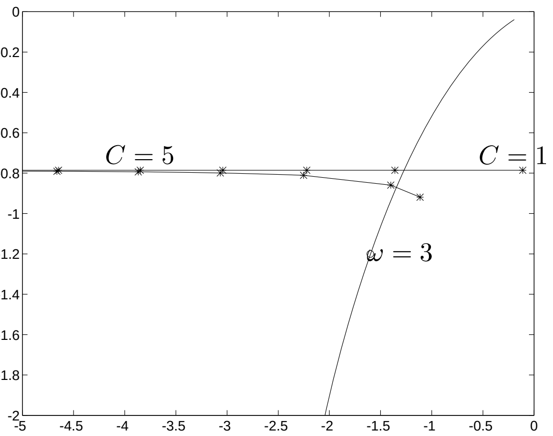


Figure 8.6: $-1/Y_1$ as a function of C for $B = 0$ (straight line) and $B = \pm 0.5$ (curved line), together with the Nyquist curve of $1/(s+1)^2$

is assumed to consist of two equal time constants. To get a symmetric relay we can rewrite the system with an external signal as in Figure 8.5. We assume that the reference signal is constant $r = r_0$, but because of the relay we assume that an oscillation exists in the loop. As usual we hope that harmonics will be damped by the linear system. We therefore try

$$e = B + C\sin\omega t = B + \frac{C}{2i}e^{i\omega t} - \frac{C}{2i}e^{-i\omega t}$$

We then get, tracing the signals around the loop.

$$B = r_0 - G(0)(P + Y_0(B,C)B)$$
$$\frac{C}{2i} = 0 - G(i\omega)\,Y_1(B,C)\,\frac{C}{2i}$$

We thus get the following system of equations

$$\begin{aligned} B + Y_0(B,C)G(0)B &= r_0 - G(0)P \\ Y_1(B,C)G(i\omega) &= -1 \end{aligned} \tag{8.27}$$

Solving this system of equations (if possible) we will get the approximate average level, as well as oscillation amplitude and frequency. The second equation of (8.27) can be interpreted as the intersection of $-1/Y_1$ and the Nyquist curve $G(i\omega)$. This is shown in Figure 8.6 for the case $P = 1$, $D = 1$. We see that the frequency and amplitude of the oscillation do not vary drastically for reasonable changes in B. Since

$$Y_0 = \frac{1}{\pi B}\left(\arcsin\left(\frac{1+B}{C}\right) - \arcsin\left(\frac{1-B}{C}\right)\right) \approx \frac{2}{\pi\sqrt{C^2-1}}$$

for small values of B and $C \approx 1.8$ we get $Y_0 \approx 0.4$, and consequently

$$B \approx 0.1(r_0 - 20), \quad C \approx 1.8, \quad \omega \approx 3.5 \tag{8.28}$$

for values of r_0 that are not too far from 20. ■

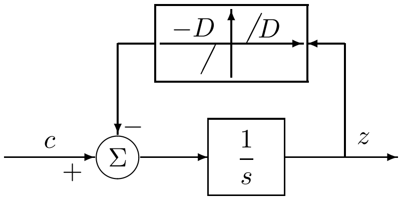


Figure 8.7: Limited integrator

8.3 Some design ideas based on the describing function

The describing function can be used to design regulators in a manner similar to classical Bode-Nyquist-Nichols techniques. This is because the describing function of the nonlinearity is treated as a transfer function (although amplitude dependent). Looking at the describing function as a parameter that lies between certain bounds, it is also possible to use robust design methods. We will here look at some examples where the describing function has inspired nonlinear compensation techniques.

8.3.1 Modified integrators

Most controllers contain integral action. The integrator gives high gain at low frequencies at the cost of a negative phase shift. In applications there is a limit to the signal level from the integrator which is needed or useful. If the output of the integrator is run through a saturation, the negative phase shift remains. Therefor a circuit like the one in Figure 8.7 is sometimes used. If the gain of the linear part of the dead zone is high enough, the output of the integrator will never rise much above ±D. If the input is

$$c(t) = C \sin \omega t$$

the output will then be

$$z(t) = \begin{cases} \dfrac{C}{\omega}(1 - \cos\theta) - D & 0 \le \theta \le \theta_0 \\ \\ D & \theta_0 < \theta \end{cases} \tag{8.29}$$

where $\theta = \omega t$ and θ_0 is given by

$$\cos\theta_0 = 1 - \frac{2\omega D}{C}$$

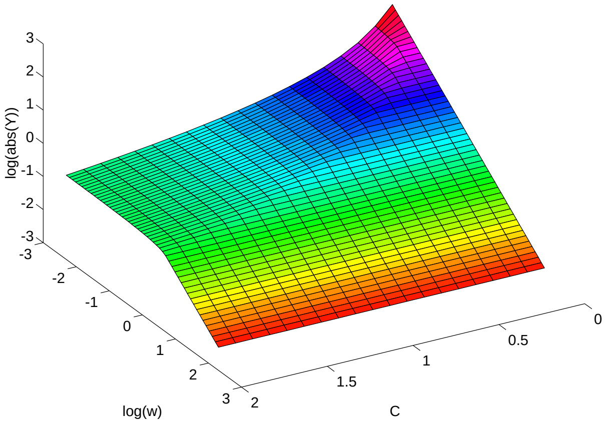


Figure 8.8: $\log|Y|$ as a function of C and $\log(\omega)$ for a limited integrator.

The expression (8.29) is valid if

$$C \geq \omega D$$

If $C < \omega D$ there is no limitation and the output is simply

$$z(t) = -\frac{C}{\omega}\cos\theta$$

A calculation of the Fourier coefficients of (8.29) gives the describing function

$$\begin{aligned} Y(C,\omega) & \\ &= -i\frac{1}{\omega}, \quad \text{if } \beta < D \\ &= \frac{1}{2\pi C}\left(4(2D-\beta)\cos\theta_0 + 3\beta + \beta\cos 2\theta_0 - \right. \\ &\qquad \left. i\left(4(2D-\beta)\sin\theta_0 + 2\beta\theta_0 + \beta\sin 2\theta_0\right)\right) \quad \text{if } \beta \geq D \end{aligned} \tag{8.30}$$

where $\beta = C/\omega$. The amplitude and phase diagrams of the describing function (8.30) are shown in Figures 8.8 and 8.9 for $D = 1$.

8.3.2 Dither techniques

Looking at Example 8.5, equation (8.28), we see that the control system actually acts like a linear system in B for small values of the bias. This is despite the fact that the relay is a highly nonlinear component. The reason is the presence of the oscillation with amplitude C. Its effect is to produce an average of the relay gain around the hysteresis loop, which turns out to be almost linear for the bias component. In the temperature control system the oscillation was produced as a by-product of the relay action in the control loop. However it is also possible to achieve similar effects by introducing external signals. Consider the following situation

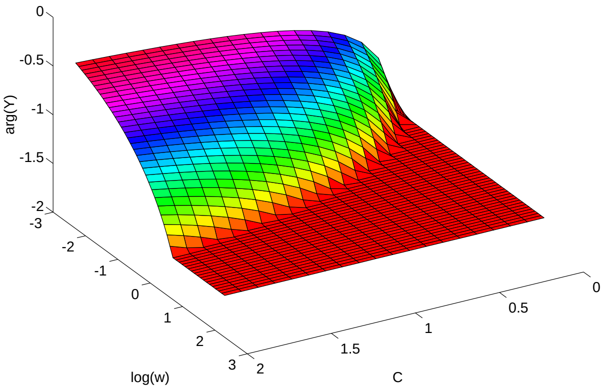
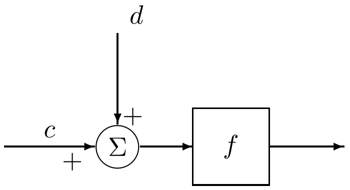


Figure 8.9: $arg(Y)$ as a function of C and $\log(\omega)$ for a limited integrator.



The signal d is periodic with a frequency which is high compared to the frequency content of c. If d is a sine:

$$d(t) = C \sin \omega t$$

and c is slowly varying compared with d, then we can regard c as a constant and use the analysis of section 8.1.2. The output will then be of the form

$$z(t) = Y_0(c(t), C)c(t) + w(t)$$

where Y_0 is a describing function of the nonlinearity f. The signal w will typically contain the frequencies ω, 2ω,... If the rest of the system acts as a low pass filter, those frequencies will dissappear and the signal d has in fact changed the nonlinearity from

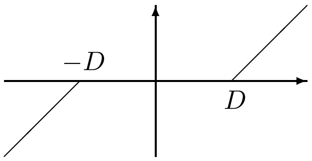
$$z(t) = f(c(t))$$

to

$$z(t) = Y_0(c(t), C)c(t)$$

The signal d is often called a *dither* signal.

Example 8.6 Let the nonlinearity be a dead zone



If the slopes of the linear parts are 1, then (8.24) gives

$$Y_0 = 1 + \frac{C}{\pi B}\left(\left(\sqrt{1-\beta_-^2} - \sqrt{1-\beta_+^2}\right) - \beta_+ \arcsin \beta_+ + \beta_- \arcsin \beta_-\right)$$

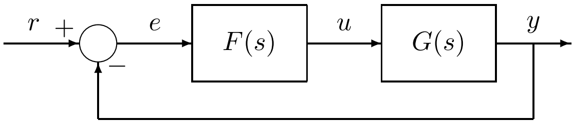
where $\beta_- = (B - D)/C$ and $\beta_+ = (B + D)/C$. If C is large enough, then

$$Y_0 \approx 1 - \frac{2D}{\pi C}$$

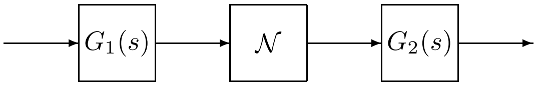
which is independent of B, so that the nonlinearity is linearized by the dither signal. ∎

8.3.3 Compensation of nonlinearities

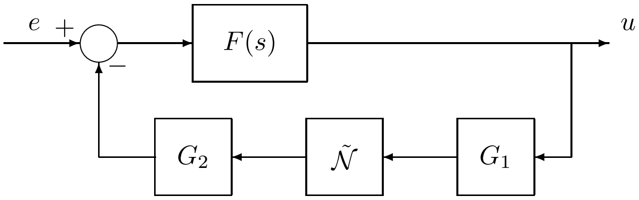
Consider a linear control system



where F is the controller and G is the controlled physical system. Suppose that the physical system actually contains a nonlinearity and is described by



where $G_1 G_2 = G$. An old idea for compensating the nonlinearity is to feed back a model of the system inside the regulator. The regulator structure then becomes



Ideally the nonlinearity $\tilde{\mathcal{N}}$ is such that

$$\mathcal{N} + \tilde{\mathcal{N}} = 1 \tag{8.31}$$

Assume that there is no bias and that higher harmonics are sufficiently damped by the linear parts. Then the nonlinearities $\tilde{\mathcal{N}}$ and $\mathcal{N}$ can be replaced by single describing functions $\tilde{Y}$ and Y respectively. The regulator is then given by

$$\tilde{F} = \frac{F}{1 + FG\tilde{Y}}$$

and the transfer function from r to y becomes

$$G_c = \frac{\tilde{F}GY}{1 + \tilde{F}GY} = \frac{FGY}{1 + FG(Y + \tilde{Y})} \tag{8.32}$$

If it is possible to achieve (8.31), then $Y + \tilde{Y} = 1$ and we get

$$G_c = \frac{FG}{1 + FG}\ Y$$

which is the closed loop transfer function of the linear system in series with Y. In particular the stability properties (determined by $1 + FG$) are not affected by the nonlinearity in this ideal case. In the more realistic case when it is not possible to achieve (8.31) exactly, the effects can be analyzed using (8.32).

8.4 Accuracy of the describing function method

Since the describing function method is based on an approximation there is always some uncertainty concerning conclusions that are based on it. There are some possibilities of being precise about the uncertainty by estimating the influence of the higher order harmonics. Consider the situation described in figure 8.10. There is no external input and we use the describing function to try and predict the presence of a self-sustained oscillation. Assuming

$$e = C \sin \omega t \tag{8.33}$$

and following the signals around the loop we get

$$G(i\omega)Y_f(C) = -1 \tag{8.34}$$

where Y_f is the describing function of f, computed from one of the formulas (8.12) – (8.14). In reality e would have the form

$$e = C \sin \omega t + e_h \tag{8.35}$$

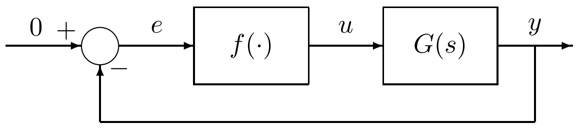


Figure 8.10: System structure for estimation of describing function accuracy.

where e_h contains the higher harmonics. Our approximation lies in neglecting e_h when computing Y_f. We will analyze the approximation under the following assumptions.

1. The nonlinearity f is odd:
$$f(-e) = -f(e) \tag{8.36}$$

2. The nonlinearity f is single-valued and lies between lines with slopes α and β:
$$\alpha e^2 \leq ef(e) \leq \beta e^2 \tag{8.37}$$

3. An oscillation has only odd harmonics. (This is natural considering the conditions on f.)

From the condition (8.37) on f it is easy to see that the describing function satisfies
$$\alpha \leq Y_f(C) \leq \beta \tag{8.38}$$
We define the average gain, f_0, and the deviation from the average gain, r:
$$f_0 = \frac{\alpha + \beta}{2}, \quad r = \frac{\beta - \alpha}{2} \tag{8.39}$$
Since our intuitive motivation for the describing function technique is based on the assumption that the harmonics are attenuated by G, it is natural to try to estimate the extent of the attenuation. This can be done by calculating the gains
$$G(3i\omega),\ G(5i\omega),\ G(7i\omega), \ldots$$
for a given ω) (remember that we only consider oscillations with odd harmonics). Define the quantities
$$\rho_k(\omega) = \left| \frac{1}{G(ki\omega)} + f_0 \right|, \quad k = 3, 5, 7, \ldots \tag{8.40}$$
$$\rho(\omega) = \inf(\rho_3(\omega), \rho_5(\omega), \ldots) \tag{8.41}$$
For those values of ω for which $\rho(\omega) > r$ we define
$$\sigma(\omega) = \frac{r^2}{\rho(\omega) - r} \tag{8.42}$$
It is now possible to state conditions for *not* having an oscillation.

Theorem 8.1 Consider a system with the structure given by Figure 8.10 whose nonlinearity satisfies (8.36) and (8.37). If

$$\left|\frac{1}{G(ki\omega)} + f_0\right| > r, \quad k = 1, 3, 5, \dots \tag{8.43}$$

then there is no periodic solution with fundamental frequency ω and only odd harmonics.
If the distance from every point of $-Y_f(C)$ to $1/G(i\omega)$ is greater than $\sigma(\omega)$, then there is no periodic solution with fundamental frequency ω and only odd harmonics.

Proof. See Khalil, [**?**]. ■

The conditions have a graphical interpretation. Equation (8.34) can be rewritten

$$\frac{1}{G(i\omega)} + Y_f(C) = 0 \tag{8.44}$$

so the describing function solution is obtained as the crossing of the inverse Nyquist plot $1/G(i\omega)$ by the describing function locus $-Y_f(C)$. For the class of functions we are considering the locus lies on the negative real axis between $-\alpha$ and $-\beta$. Figure 8.11 shows the two situations where no oscillation can take place for a certain ω. To get conditions that guarantee an oscillation to take place, one has to look at the uncertainty band around $1/G(i\omega)$ defined by the circles of radius $\sigma(\omega)$. It is the possible intersection of this band with the $-Y_f$-locus that is important. The geometrical situation is shown in Figure 8.12. Let C_1 and C_2 be the amplitudes for which the uncertainty bounds intersect $-Y_f(C)$. Let ω_1 and ω_2 be the frequencies corresponding to uncertainty circles which are tangent to $-Y_f$ from above and below. Together they define an uncertainty rectangle in the amplitude-frequency space:

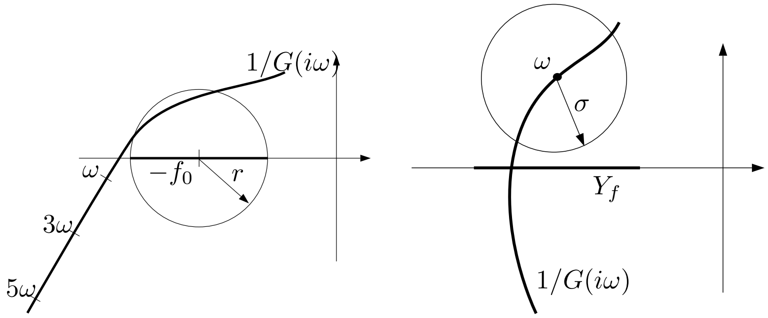


Figure 8.11: Two situations when no oscillation can take place for a certain ω.

$$\Gamma = \{(\omega, C) : \omega_1 < \omega < \omega_2,\ C_1 < C < C_2\} \tag{8.45}$$

We also define ω_o and C_o to be the frequency and amplitude for the nominal intersection between $1/G$ and $-Y$, i.e. the values that the describing function method would give us. The basic result is then as follows.

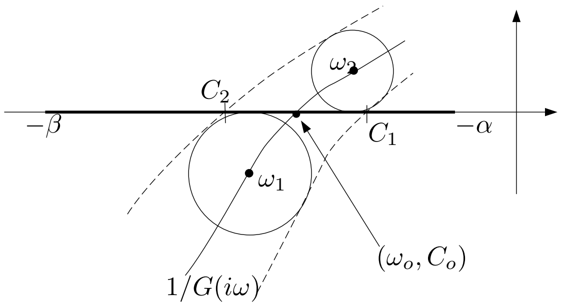


Figure 8.12: Intersection of $1/G$ and its uncertainty band with the locus of $-Y_f$.

Theorem 8.2 Consider a system with the structure given by Figure 8.10 whose nonlinearity satisfies (8.36) and (8.37). Let there be an uncertainty band which intersects the locus of $-Y_f(C)$ in such a way that the set Γ of (8.45) is well defined. Let there be a unique intersection of $1/G(i\omega)$ and $-Y_f(C)$ in Γ at ω_o, C_o. Assume that

$$\frac{dY_f}{dC}(C_o) \neq 0, \qquad \frac{d\mathrm{Im}G(i\omega)}{d\omega}(i\omega_o) \neq 0$$

Then there exists a periodic solution, having odd harmonics only, of the form

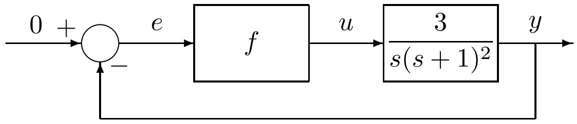
$$e(t) = C \sin \omega t + e_h(t)$$

where ω, C belong to $\bar{\Gamma}$ and

$$\frac{\omega}{\pi} \int_0^{2\pi/\omega} e_h(t)^2 \, dt \leq \left(\frac{\sigma(\omega) C}{r} \right)^2$$

Proof. See Khalil, [**?**]. ∎

Example 8.7 Consider the system



where f is a saturation:

$$f(e) = \begin{cases} e & |e| \leq 1 \\ 1 & e > 1 \\ -1 & e < -1 \end{cases}$$

The describing function is

$$Y_f(C) = \begin{cases} \frac{2}{\pi}(\arcsin \frac{1}{C} + \frac{1}{C}\sqrt{1 - C^{-2}}) & C > 1 \\ 1 & C \leq 1 \end{cases}$$

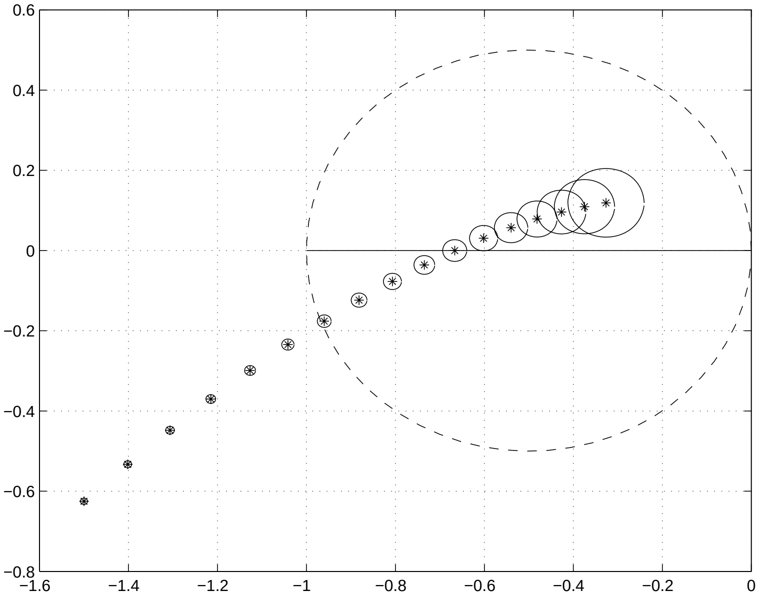


Figure 8.13: $1/G(i\omega)$ for $\omega = 0.7, 0.75, 0.8, \ldots, 1.5$ with error circles. The solid line is $-Y_f$.

In Figure 8.13 the inverse Nyquist curve is plotted together with $-Y_f$ and some error circles. The nominal solution is $\omega_o = 1$ and $C_o = 1.8$. We see that $\omega_1 \approx 0.95$ and that ω_2 is slightly less than 1.05. The uncertainty bounds intersect the real axis at -0.59 and -0.72 corresponding to $C_1 = 1.6$ and $C_2 = 2.1$. We conclude that there exists a periodic solution whose fundamental component has

$$0.95 \leq \omega \leq 1.05, \quad 1.6 \leq C \leq 2.1$$

The nominal solution is $\omega_o = 1$, $C_o = 1.8$. A simulation is shown in Figure 8.14.

■

8.5 Exercises

8.1 Compute the describing functions for bias plus sinusoid for a piecewise linear system:

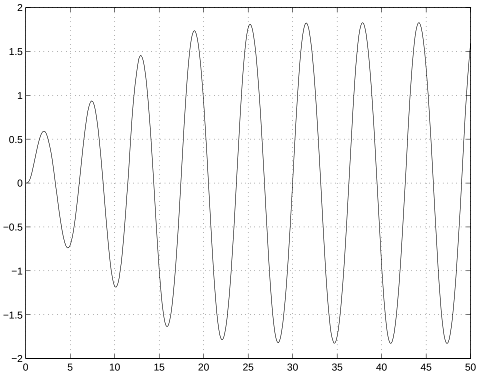
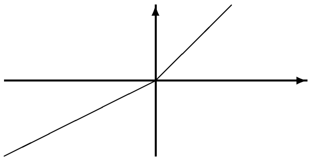
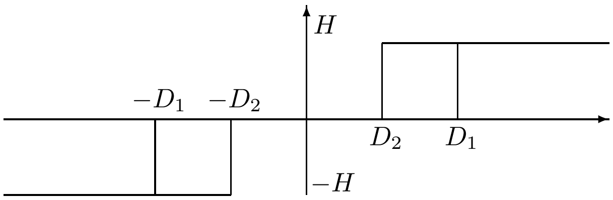


Figure 8.14: Simulation of periodic solution.



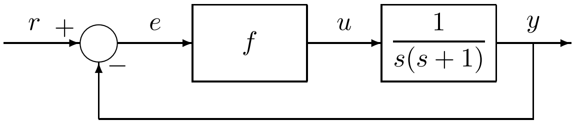
Assume the slopes to be k_1 and k_2 respectively.

8.2 Compute the describing function for a sinusoid without bias when the nonlinearity is a relay with dead zone and hysteresis:

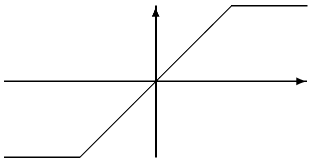


8.3 Consider the temperature control system of Example 8.5. Suppose the nonlinearity is a relay with dead zone and hysteresis as described in the previous exercise. In which way is the system behavior altered? What happens if additional time constants are introduced in the linear system?

8.4 Consider the system

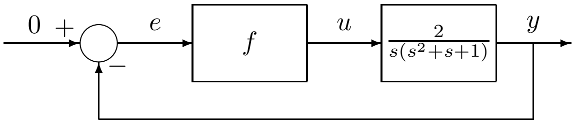


where f is a saturation:



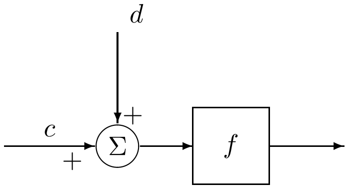
Let r be a sinusoid without bias. Compute the gain from r to e as a function of frequency and amplitude. What happens if r has a bias?

8.5 Consider the system



where f is a saturation. Compute the amplitude and frequency of any periodic solution together with error bounds.

8.6 Consider a nonlinearity with dither signal:



Suppose the dither signal is a sawtooth signal and the nonlinearity an ideal relay (that is without dead zone or hysteresis). Show that the effective nonlinearity is a saturation. How do the parameters of the saturation depend on the parameters of the sawtooth signal?

.

Chapter 9

Tracking and disturbance rejection

To follow a reference signal and reject disturbances is typically the main task of a control system. When nonlinearities are present this requires some extra care.

9.1 A simple nonlinear servo problem

Consider the simple system of Figure 9.1. If we assume the reference signal to be constant, the control system is described by the following equations

$$\dot{x} = -x^3 + Ke \tag{9.1}$$

$$\dot{r} = 0 \tag{9.2}$$

$$e = r - x \tag{9.3}$$

We see that the description contains three parts, a model of the system with controller (9.1), a model for the reference signal (9.2) and a description of the error signal (9.3). We would like the error to be zero, at least in the steady state. However we get

$$(r - e)^3 = Ke \tag{9.4}$$

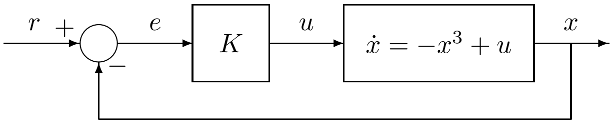


Figure 9.1: Control system with P-controller

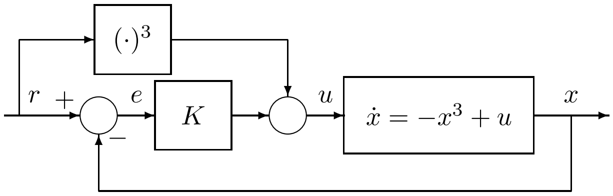


Figure 9.2: Control system with P-controller and feed-forward.

for the steady state. If K is large enough we get

$$e \approx \frac{r^3}{K} \tag{9.5}$$

It is not surprising that we get a steady state error, since that is what we would get also in the linear case. The nonlinearity makes itself felt from the fact that the error is proportional to the reference signal cubed, rather than the reference signal itself. If we want to remove the steady state error, we can introduce a feed-forward compensation as in figure 9.2. The system equations are now

$$\dot{x} = -x^3 + Ke + r^3 \tag{9.6}$$

$$\dot{r} = 0 \tag{9.7}$$

$$e = r - x \tag{9.8}$$

In equilibrium ($\dot{x} = 0$) we have

$$0 = -(r-e)^3 + Ke + r^3 \tag{9.9}$$

which reduces to

$$(K + 3r^2)e - 3re^2 + e^3 = 0 \tag{9.10}$$

The only solution is $e = 0$, so the tracking error is zero in steady state.

9.2 General problem formulation

Let us now place the problem mentioned in the preceding section in a general framework. Suppose we have a control problem where there are both reference signals to track and disturbance signals to reject. We can group these signals together in a vector $v(t)$ (dimension q). We assume that we have a model of the reference and disturbance signals:

$$\dot{w} = g(w) \tag{9.11}$$

This model is often referred to as an *exosystem.* (In our example of the previous section the model was just $\dot{r} = 0$.) The system to be controlled is given by a state space description

$$\dot{x} = f(x, u, w) \tag{9.12}$$

with input u (m-vector) and state x (n-vector). We assume that the control objective can be expressed by an error signal e (p-vector):

$$e = h(x, w) \tag{9.13}$$

The total system description is then of the form

$$\begin{aligned} \dot{x} &= f(x, u, w) \\ \dot{w} &= g(w) \\ e &= h(x, w) \end{aligned} \tag{9.14}$$

Compare this with (9.1), (9.2), (9.3). To simplify calculations, we assume all the functions f, g and h to be infinitely differentiable. In addition we postulate that

$$f(0, 0, 0) = 0, \quad g(0) = 0, \quad h(0, 0) = 0 \tag{9.15}$$

that is, the origin is an equilibrium point of the system for $u = 0$.

We want a control configuration that achieves the following objectives, the *Global Tracking Problem.*

- The closed loop system is globally exponentially stable when $w = 0$.
- The error goes to zero:

$$\lim_{t\to\infty} e(t) = 0, \quad \text{for all } x(0), w(0) \tag{9.16}$$

However we will have to relax the requirement and be content with solutions that are guaranteed to work locally. There are many possible controller configurations, but we will look at the following two typical cases.
Pure state feedback

$$u = k(x, w) \tag{9.17}$$

Note that the feedback is from all states, including w. This means that we are actually using a controller that has *feed-forward* from the reference and disturbance signals.
Pure error feedback

$$u = k(\xi), \quad \dot{\xi} = m(\xi, e) \tag{9.18}$$

Here we assume that the error is the input to a dynamic controller, which is the classical control configuration.

What assumptions should we make about the exosystem? In linear control theory one often assumes that the exosystem has all its eigenvalues on the imaginary axis or in the right half plane. This is because modes of the exosystem that decay to zero are trivial, since they do not contribute to the error as time goes to infinity. On the other hand it is natural to allow instability, since one might be interested in tracking signals that are growing, like ramps. For nonlinear systems however, unbounded signals are difficult to handle, so we will assume that the exosystem is stable. To ensure that the exosystem is difficult enough to require some kind of control action asymptotically, we will introduce the following notion.

Definition 9.1 Consider a dynamic system

$$\dot{w} = g(w), \quad w(0) = w_0 \tag{9.19}$$

with solution $w(t) = \gamma(t, w_0)$. An initial value w_0 is said to be *Poisson stable* if, for every neighborhood U of w_0 and every $T > 0$ there exist $t_1 > T$ and $t_2 < -T$ such that $\gamma(t_1, w_0) \in U$ and $\gamma(t_2, w_0) \in U$.

Poisson stability of a state thus means that the solution comes back "almost to the same point", infinitely many times. Summarizing, we make the following assumption.

Assumption I. The origin is a stable equilibrium of the exosystem, and there is a neighborhood of the origin, where every w is Poisson stable.

A system satisfying Assumption I is sometimes said to be *neutrally stable*. Let us now try to solve the problem for the feedback case, by looking at the linearizations at the origin. The local description of the system is

$$\begin{aligned} \dot{x} &= Ax + Bu + Fw + \phi_1(x, u, w) \\ \dot{w} &= Gw + \psi(w) \\ e &= Cx + Dw + \phi_2(x, w) \end{aligned} \tag{9.20}$$

where ϕ_1, ϕ_2 and ψ are functions that vanish at the origin, together with their first order derivatives, and

$$A = f_x(0,0,0), \quad B = f_u(0,0,0), \quad F = f_w(0,0,0), \quad G = g_w(0) \tag{9.21}$$

$$C = h_x(0,0), \quad D = h_w(0,0) \tag{9.22}$$

If we assume a feedback

$$u = k(x, w) = Kx + Lw + \phi_3(x, w) \tag{9.23}$$

then the dynamics of the closed loop system becomes

$$\frac{d}{dt}\begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} A + BK & BL + F \\ 0 & G \end{pmatrix}\begin{pmatrix} x \\ w \end{pmatrix} + \phi(x, w) \tag{9.24}$$

where ϕ is a function that vanishes at the origin together with its Jacobian. Let us consider the eigenvalues of the linear part. If the linearized system is stabilizable we can choose K so that $A + BK$ has all its eigenvalues in the left half plane. The stability of the exosystem implies that G can have no eigenvalues with positive real part. The Poisson stability means that there can be no eigenvalues with negative real part. The matrix G thus has all its eigenvalues on the imaginary axis. Systems with the property that the linearization has some eigenvalues on the imaginary axis have received considerable intention and there exists a collection of results, known as "center manifold theory". We will consider it in the next section.

9.3 Center manifold theory

Consider a system

$$\dot{x} = Ax + f(x), \quad f(0) = 0, \quad f_x(0) = 0 \tag{9.25}$$

where x is an n-vector and f is a k times ($k \geq 2$) continuously differentiable function. The linearization is then

$$\dot{y} = Ay \tag{9.26}$$

Let E_s, E_u and E_c be the vector spaces spanned by eigenvectors of A corresponding to eigenvalues with negative, positive and zero real parts respectively. The space E_s thus contains solutions of (9.26) converging to the origin, while E_u contains unstable solutions of (9.26) that start at the origin. We expect solutions of (9.25) to be in some sense similar to those of the linearization (9.26), at least close to the origin. One can formalize this by looking at *invariant manifolds*. A manifold is a subset that "locally looks like" an open subset of $\mathbb{R}^d$ for some value of d. A manifold is invariant if solutions of (9.25) that start in the manifold remain there, when going backwards or forwards in time. Typically a manifold is described by an equation

$$g(x) = 0 \tag{9.27}$$

where g is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ whose Jacobian is nonsingular. The dimension of the manifold will then be $n - m$. With this terminology we can describe the relation between (9.25) and (9.26) as follows.

Theorem 9.1 For the system (9.25) there exist invariant manifolds W_s ("the stable manifold"), W_u ("the unstable manifold") and W_c ("the center manifold") that are tangent to E_s, E_u and E_c respectively at the origin.

In our control theory application of tracking and disturbance rejection we are only interested in systems without any unstable manifold. If we assume the linear part to be block diagonal we can write

$$\begin{aligned} \dot{y} &= Gy + g(y, z) \\ \dot{z} &= Hz + h(y, z) \end{aligned} \tag{9.28}$$

where G is a matrix having eigenvalues with strictly negative real parts, and H is a matrix whose eigenvalues all lie on the imaginary axis. The functions g and h are zero at the origin together with their Jacobians. The center manifold can then locally be described by a function

$$y = \pi(z) \tag{9.29}$$

Since a center manifold is tangent to E_c at the origin we get

$$\pi(0) = 0, \quad \frac{\partial \pi}{\partial z}(0) = 0 \tag{9.30}$$

The invariance gives the relation $\dot{y} = \frac{\partial \pi}{\partial z}\dot{z}$. From (9.28) we then get

$$\frac{\partial \pi(z)}{\partial z}\ (Hz + h(\pi(z), z)) = G\pi(z) + g(\pi(z), z) \tag{9.31}$$

which is a partial differential equation for π from which we can in principle calculate the center manifold.

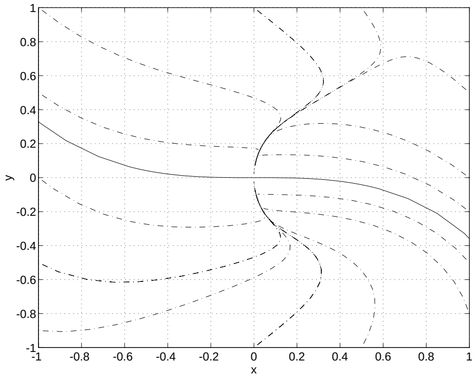


Figure 9.3: Phase portrait for (9.32). The stable and center manifolds are solid while various trajectories are dash-dotted.

Example 9.1 Consider the system

$$\begin{aligned} \dot{x} &= -x + y^2 \\ \dot{y} &= -y^3 + x^3 \end{aligned} \tag{9.32}$$

The subspaces E_s and E_c are the x-axis and y-axis respectively. In Figure 9.3 the phase plane is shown. The stable manifold is shown as a solid curve and is tangent to the x-axis as expected. All trajectories that do not start on the stable manifold tend to approach the same curve asymptotically. This curve, which is tangent to the y-axis is a center manifold. The equation for the center manifold (9.31) becomes

$$\pi'(y)(-y^3 + \pi(y)^3) = -\pi(y) + y^2 \tag{9.33}$$

From (9.30) it follows that close to the origin we have

$$\pi(y) = \alpha y^2 + \beta y^3 + O(y^4) \tag{9.34}$$

Substituting this expression into (9.33) gives

$$\pi(y) = y^2 + O(y^4) \tag{9.35}$$

The center manifold is thus locally a parabola, symmetric about the positive x-axis. Substituting the expression for π into the system equations gives

$$\dot{y} = -y^3 + y^6 + O(y^8) \tag{9.36}$$

as the dynamics on the center manifold. Solutions that start on the manifold close enough to the origin will thus converge slowly to the origin. ∎

In the example solutions seem to approach the center manifold rapidly. This behavior is typical.

Theorem 9.2 Suppose $y = \pi(z)$ is a center manifold of (9.28) at the origin. There is a neighborhood U of the origin and positive numbers K and C such that, if $y(0), z(0)$ is in U then

$$|y(t) - \pi(z(t))| \leq Ce^{-Kt}|y(0) - \pi(z(0))| \tag{9.37}$$

as long as $(y(t), z(t))$ is in U.

Definition 9.2 We call a system

$$\dot{x} = f(x), \quad f(x_0) = 0, \quad A = f_x(0) \tag{9.38}$$

exponentially stable at x_0 if all eigenvalues of A have strictly negative real parts.

9.4 State feedback

We now return to the state feedback problem. We begin by a more careful formulation of what we want to achieve. Consider the system

$$\begin{aligned} \dot{x} &= f(x, u, w) \\ \dot{w} &= g(w) \\ e &= h(x, w) \end{aligned} \tag{9.39}$$

Definition 9.3 We will look at a local version of the global tracking problem presented in section 9.2. The *local state feedback tracking problem* consists of finding a controller

$$u = k(x, w) \tag{9.40}$$

such that

1. The closed loop system

   $$\dot{x} = f(x, k(x, 0), 0) \tag{9.41}$$

   is locally (at the origin) exponentially stable.

2. All solutions of

   $$\begin{aligned} \dot{x} &= f(x, k(x, w), w) \\ \dot{w} &= g(w) \\ e &= h(x, w) \end{aligned} \tag{9.42}$$

   starting sufficiently close to the origin, are such that

   $$\lim_{t \to \infty} e(t) = 0 \tag{9.43}$$

Now consider (9.24). We realize that there must exist a center manifold

$$x = \pi(w) \tag{9.44}$$

Since solutions starting on the center manifold remain there, we get

$$\dot{x} = \pi_w \dot{w} \tag{9.45}$$

Substituting the differential equations satisfied by x and w, we get

$$\pi_w(w) \; g(w) = f(\pi(w), k(\pi(w), w), w) \tag{9.46}$$

We can then show the following result.

Theorem 9.3 If the local state feedback tracking problem is solvable, then the pair A, B is stabilizable and there exist functions $c(w)$ and $\pi(w)$ satisfying

$$\pi_w(w) \; g(w) = f(\pi(w), c(w), w) \tag{9.47}$$

$$h(\pi(w), w) = 0 \tag{9.48}$$

Proof. Equation (9.47) was shown in (9.46). To show (9.48) consider a point $(w_0, \pi(w_0)$ on the center manifold. For any $\epsilon > 0$ and any $T > 0$ we can, by Assumption I, find a $t > T$ such that $|w(t) - w_0| < \epsilon$. Since π is smooth, we can actually find a $t > T$ such that

$$|w(t) - w_0| < \epsilon, \quad |\pi(w(t)) - \pi(w_0)| < \epsilon \tag{9.49}$$

For any point on the center manifold the solution will thus come back arbitrarily close to that point infinitely many times. The only way of achieving the requirement $e(t) \to 0$ is then to have the error equal to zero on the center manifold, that is

$$h(\pi(w), w) = 0 \tag{9.50}$$

To get exponential stability we see from (9.24) that there has to be a K such that $A+BK$ has all its eigenvalues strictly in the left half plane. This is precisely the definition of stabilizability for the pair A, B. ∎

The conditions of the theorem are also sufficient as shown by the construction of the following theorem.

Theorem 9.4 Suppose there exist continuously differentiable functions π and c, with $\pi(0) = 0$, $c(0) = 0$ such that (9.47), (9.48) hold and assume that A, B is stabilizable. Choose a function $\bar{k}(x)$ such that $A + B\bar{k}_x(0)$ has all its eigenvalues strictly in the left half plane. Then the control law

$$u = k(x, w) = c(w) + \bar{k}(x - \pi(w)) \tag{9.51}$$

solves the local state feedback tracking problem.

Proof. The first property of Definition 9.3 is trivially satisfied since

$$k(x, 0) = c(0) + \bar{k}(x - \pi(0)) = \bar{k}(x) \tag{9.52}$$

The linearization at the origin of $f(x, k(x, 0), 0)$ is $(A + B\bar{k}_x(0)$, showing the exponential stability. The closed loop system will then be given by

$$\begin{aligned} \dot{x} &= f(x, c(w) + \bar{k}(x - \pi(w)), w) \\ \dot{w} &= g(w) \end{aligned} \tag{9.53}$$

The right hand side has the linearization

$$\begin{pmatrix} A + B\bar{k}_x(0) & * \\ 0 & G \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} \tag{9.54}$$

Replacing x with $\pi(w)$ in $f(x, c(w) + \bar{k}(x - \pi(w)), w)$ gives $f(\pi(w), c(w), w)$, so the center manifold equation for (9.53) is precisely the first equation of (9.47). Equation (9.48) shows that the error is zero on the center manifold. Since Theorem 9.2 shows exponential convergence to the center manifold we conclude that the error decays exponentially, so property 2 of Definition 9.3 is also satisfied. ■

Example 9.2 Consider the simple control system of Section 9.1. The system dynamics without the controller is

$$\begin{aligned} \dot{x} &= -x^3 + u \\ \dot{r} &= 0 \\ e &= r - x \end{aligned}$$

The center manifold equations are

$$0 = -\pi(r)^3 + c(r), \quad 0 = r - \pi(r) \tag{9.55}$$

giving

$$\pi(r) = r, \quad c(r) = r^3 \tag{9.56}$$

Modifying the P-controller according to (9.51) gives

$$u = r^3 + K(r - x) \tag{9.57}$$

which is precisely the compensation shown in Figure 9.2. ■

9.5 Error feedback

We will now define the error feedback problem for the system

$$\begin{aligned} \dot{x} &= f(x, u, w) \\ \dot{w} &= g(w) \\ e &= h(x, w) \end{aligned} \tag{9.58}$$

more precisely. We will assume that the regulator is a general nonlinear system with state ξ, input e and output u:

$$\begin{aligned} u &= k(\xi) \\ \dot{\xi} &= m(\xi, e) \end{aligned} \tag{9.59}$$

We then have the following problem to solve.

Definition 9.4 The *local error feedback tracking problem* consists of finding a controller

$$\begin{aligned} u &= k(\xi) \\ \dot{\xi} &= m(\xi, e) \end{aligned} \tag{9.60}$$

such that

1. The closed loop system

$$\begin{aligned}\dot{x} &= f(x, k(\xi), 0)\\ \dot{\xi} &= m(\xi, h(x, 0))\end{aligned} \tag{9.61}$$

   is exponentially stable at the origin.

2. All solutions of

$$\begin{aligned}\dot{x} &= f(x, k(\xi), w)\\ \dot{w} &= g(w)\\ \dot{\xi} &= m(\xi, h(x, w))\\ e &= h(x, w)\end{aligned} \tag{9.62}$$

   starting sufficiently close to the origin, are such that

$$\lim_{t\to\infty} e(t) = 0 \tag{9.63}$$

To formulate the conditions for solving the local error feedback tracking problem it is useful to have the following concept.

Definition 9.5 The system

$$\dot{x} = f(x) \tag{9.64}$$

$$y = h(x) \tag{9.65}$$

is *immersed* into the system

$$\dot{\bar{x}} = \bar{f}(\bar{x}) \tag{9.66}$$

$$y = \bar{h}(\bar{x}) \tag{9.67}$$

if there exists a smooth mapping $\bar{x} = \tau(x)$ such that

$$\tau_x f(x) = \bar{f}(\tau(x)) \tag{9.68}$$

$$h(x) = \bar{h}(\tau(x)) \tag{9.69}$$

for all x.

Equation (9.68) in the definition is another way of saying that $\frac{d}{dt}\tau(x) = \bar{f}(\tau(x))$, for a solution to (9.64) i.e. that solutions to the x-system are carried into solutions to the $\bar{x}$-system. Equation (9.69) shows that this is done in such a way that the output is the same. The definition thus says that any output that can be produced by (9.64), (9.65) can also be produced by (9.66), (9.67) by taking $\bar{x}(0) = \tau(x(0))$.

We can now state the basic result.

Theorem 9.5 Consider the local error feedback tracking problem with assumption I satisfied. It is solvable if and only if there exist functions $\pi(w)$ and $c(w)$ with $\pi(0) = 0$, $c(0) = 0$ such that

$$\begin{aligned}\pi_w(w)\, g(w) &= f(\pi(w), c(w), w)\\ h(\pi(w), w) &= 0\end{aligned} \tag{9.70}$$

and such that the system

$$\begin{aligned}\dot{w} &= g(w)\\ u &= c(w)\end{aligned} \tag{9.71}$$

with output u is immersed into a system

$$\begin{aligned}\dot{\xi} &= \phi(\xi)\\ u &= \gamma(\xi)\end{aligned} \tag{9.72}$$

in which $\phi(0) = 0$, $\gamma(0) = 0$, $M = \phi_\xi(0)$, $K = \gamma_\xi(0)$ and where the pair

$$\begin{bmatrix} A & 0\\ NC & M\end{bmatrix}, \quad \begin{bmatrix} B\\ 0\end{bmatrix} \tag{9.73}$$

is stabilizable for some N and the pair

$$\begin{bmatrix} A & BK\\ 0 & M\end{bmatrix}, \quad \begin{bmatrix} C & 0\end{bmatrix} \tag{9.74}$$

detectable.

Proof. *Necessity.* If the problem is solvable then there exists a controller

$$\dot{\xi} = m(\xi, e), \quad u = k(\xi) \tag{9.75}$$

such that for solutions of (9.62) the error goes to zero at least locally at the origin. Define $M = m_\xi(0,0)$, $N = m_e(0,0)$, $K = k_\xi(0)$. Linearizing (9.62) at the origin we get

$$\frac{d}{dt}\begin{bmatrix} x\\ \xi\\ w\end{bmatrix} = \begin{bmatrix} A & BK & F\\ NC & M & ND\\ 0 & 0 & G\end{bmatrix}\begin{bmatrix} x\\ \xi\\ w\end{bmatrix} + \rho(x,\xi,w) \tag{9.76}$$

where ρ is a remainder term which is zero at the origin together with its Jacobian. From the center manifold theorem it follows analogously to the proof of Theorem 9.3 that there exist mappings $x = \pi(w)$, $\xi = \sigma(w)$ such that

$$\pi_w g(w) = f(\pi(w), k(\sigma(w)), w) \tag{9.77}$$

$$\sigma_w g(w) = m(\sigma(w), 0) \tag{9.78}$$

Defining $c(w) = k(\sigma(w))$ we get the first equation of (9.70) from (9.77). The second equation of (9.70) is shown using the same argument as in the proof of Theorem 9.3. Defining $\phi(\xi) = m(\xi, 0)$ and $\gamma(\xi) = k(\xi)$ is is seen from (9.78) that (9.71) is immersed into (9.72).

From (9.76) it follows that the matrix

$$\begin{bmatrix} A & BK\\ NC & M\end{bmatrix}$$

has all its eigenvalues strictly in the left half plane. Since

$$\begin{bmatrix} A & BK\\ NC & M\end{bmatrix} = \begin{bmatrix} A & 0\\ NC & M\end{bmatrix} + \begin{bmatrix} B\\ 0\end{bmatrix}\begin{bmatrix} K & 0\end{bmatrix}$$

it follows that

$$\begin{bmatrix} A & 0 \\ NC & M \end{bmatrix}, \quad \begin{bmatrix} B \\ 0 \end{bmatrix}$$

is stabilizable and since

$$\begin{bmatrix} A & BK \\ NC & M \end{bmatrix} = \begin{bmatrix} A & BK \\ 0 & M \end{bmatrix} + \begin{bmatrix} 0 \\ N \end{bmatrix} \begin{bmatrix} C & 0 \end{bmatrix}$$

it follows that

$$\begin{bmatrix} A & BK \\ 0 & M \end{bmatrix}, \quad \begin{bmatrix} C & 0 \end{bmatrix}$$

is detectable.

Sufficiency. Since

$$\begin{bmatrix} A & 0 \\ NC & M \end{bmatrix}, \quad \begin{bmatrix} B \\ 0 \end{bmatrix}$$

is stabilizable it follows that

$$\begin{bmatrix} A & BK \\ NC & M \end{bmatrix}, \quad \begin{bmatrix} B \\ 0 \end{bmatrix}$$

is stabilizable, and since

$$\begin{bmatrix} A & BK \\ 0 & M \end{bmatrix}, \quad \begin{bmatrix} C & 0 \end{bmatrix}$$

is detectable, so is

$$\begin{bmatrix} A & BK \\ NC & M \end{bmatrix}, \quad \begin{bmatrix} C & 0 \end{bmatrix}$$

It follows that the linear system

$$\dot{z} = \begin{bmatrix} A & BK \\ NC & M \end{bmatrix} z + \begin{bmatrix} B \\ 0 \end{bmatrix} u, \quad y = \begin{bmatrix} C & 0 \end{bmatrix} x$$

is both stabilizable and detectable. Consequently there is a dynamic output feedback controller

$$\dot{\eta} = \bar{A}\eta + \bar{B}y, \quad u = \bar{C}\eta$$

which gives a closed loop system with all eigenvalues strictly in the left half plane. (The controller could for instance be based on an observer and feedback from the observer states.) The matrix

$$J = \begin{bmatrix} A & BK & B\bar{C} \\ NC & M & 0 \\ \bar{B}C & 0 & \bar{A} \end{bmatrix} \tag{9.79}$$

then has all its eigenvalues strictly in the left half plane.

Now construct the following controller

$$\begin{aligned} \dot{\xi} &= \phi(\xi) + Ne \\ \dot{\eta} &= \bar{A}\eta + \bar{B}e \\ u &= \bar{C}\eta + \gamma(\xi) \end{aligned} \tag{9.80}$$

The closed loop system is then described by

$$\frac{d}{dt}\begin{bmatrix} x \\ \xi \\ \eta \\ w \end{bmatrix} = \begin{bmatrix} A & BK & B\bar{C} & * \\ NC & M & 0 & * \\ \bar{B}C & 0 & \bar{A} & * \\ 0 & 0 & 0 & G \end{bmatrix} \begin{bmatrix} x \\ \xi \\ \eta \\ w \end{bmatrix} + \text{higher order terms}$$

The block matrix in the upper left hand corner is identical to the matrix J of (9.79) which has all its eigenvalues strictly in the left half plane. The exponential stability condition is thus satisfied. Let σ be the mapping from (9.71) to (9.72) defined by the immersion. Then $x = \pi(w)$, $\xi = \sigma(w)$ defines the center manifold. It follows from the second equation of (9.70) that the error is zero on the center manifold. From Theorem 9.2 it follows that the error goes to zero as t goes to infinity. ■

Remark 9.1 The condition that (9.71) is immersed in (9.72) means that the controller contains a model of the exosystem. This is clearly seen in the controller construction (9.80). Sometimes this fact is referred to as the *internal model principle.*

Remark 9.2 Our controller (9.80) is linear except for the terms ϕ and γ. This is sufficient to show that the tracking problem is solved *locally.* In practice one would probably want to try a nonlinear controller to extend the region of convergence.

Remark 9.3 It is easy to see that the pair (9.73) is stabilizable only when A, B. Similarly detectability of (9.74) requires detectability of A, C. A necessary condition for solving the local error feedback tracking problem is therefore that the controlled system has a linearization which is stabilizable and detectable.

Example 9.3 Consider again the simple control system equations

$$\dot{x} = -x^3 + u \tag{9.81}$$

$$\dot{r} = 0 \tag{9.82}$$

$$e = r - x \tag{9.83}$$

where we now want to use error feedback. The center manifold equations (9.70) are

$$0 = -\pi(r)^3 + c(r), \quad 0 = r - \pi(r) \tag{9.84}$$

The solution for π and c is

$$\pi(r) = r, \quad c(r) = r^3 \tag{9.85}$$

For $e = 0$, the control should thus be generated by

$$\dot{\xi} = 0, \quad u = \xi^3$$

We thus get $M = 0$, $K = 0$. It follows that the pair

$$\begin{bmatrix} A & BK \\ 0 & M \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} C & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \end{bmatrix}$$

is not detectable. The problem is thus not solvable. ■

Example 9.4 Modify the previous example so that

$$\dot{x} = -x - x^3 + u \tag{9.86}$$

$$\dot{r} = 0 \tag{9.87}$$

$$e = r - x \tag{9.88}$$

The center manifold equations (9.70) are

$$0 = -\pi(r) - \pi(r)^3 + c(r), \quad 0 = r - \pi(r) \tag{9.89}$$

The solution for π and c is

$$\pi(r) = r, \quad c(r) = r + r^3 \tag{9.90}$$

For $e = 0$, the control should thus be generated by

$$\dot{\xi} = 0, \quad u = \xi + \xi^3$$

We thus get $M = 0$, $K = 1$. If we take $N = 1$, then the matrix

$$\begin{bmatrix} A & BK \\ NC & M \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$$

has already its eigenvalues in the left half plane, so we can take the controller

$$\dot{\xi} = e, \quad u = \xi + \xi^3$$

■

9.6 Exercises

9.1 Calculate the center manifold of the system

$$\begin{aligned} \dot{x} &= -x + y + y^2 \\ \dot{y} &= x - y \end{aligned} \tag{9.91}$$

Is the dynamics on the center manifold stable or unstable?

9.2 A classical example of a system having a non-unique center manifold is

$$\begin{aligned} \dot{x} &= x^2 \\ \dot{y} &= -y \end{aligned} \tag{9.92}$$

Show that there are in fact infinitely many center manifolds and compute them all.

9.3 Consider Example 9.4 and take the linear part of the controller, i.e.

$$\dot{\xi} = e, \quad u = \xi$$

Compare the performance with the nonlinear controller, especially for large steps in the reference signal. Discuss the result.

9.4 For a linear plant with linear exosystem the conditions for solving the tracking problem are a set of linear matrix equations. What are they?

9.5 Use (9.47) to give some examples of tracking/disturbance rejection problems that are impossible to solve.

9.6 Consider the system

$$\begin{aligned}\dot{x}_1 &= u\\ \dot{x}_2 &= x_1 + v^3\end{aligned}$$

where v is a sinusoidal disturbance, whose angular frequency is known but whose amplitude and phase are unknown. It is desired to have x_2 equal to zero, at least asymptotically. Construct a controller.

Chapter 10

Physics based control

10.1 Lagrangian physics

The Lagrangian methos of modelling has its origin in mechanics but can be extended to other systems, e.g. electrical and electro-mechanical ones. The state vector is assumed to have the form

$$x = \begin{bmatrix} q \\ \dot{q} \end{bmatrix}$$

for some set of variables q. In mechanics q typically consists of distances and angles, so that $\dot{q}$ consists of velocities and angular velocities. The Langrangian modeling technique also postulates the existence of a function L of the form

$$L(q, \dot{q}) = T(q, \dot{q}) - V(q) \tag{10.1}$$

satisfying the equation

$$\frac{d}{dt} L_{\dot{q}}^T - L_q^T = Q \tag{10.2}$$

In mechanics T is the kinetic energy and V is the potential energy. We will assume that the kinetic energy has the form

$$T(q, \dot{q}) = \frac{1}{2} \dot{q}^T D(q) \dot{q}, \quad D(q) = D^T(q) > 0 \tag{10.3}$$

The vector Q is called the generalized force. In mechanics it consists of ordinary forces (corresponding to components of q that are distances) and torques (corresponding to angles in q). In control applications Q typically has the form

$$Q = -F(\dot{q}) + d + Bu \tag{10.4}$$

where F is some generalized force generated by the motion (often some type of friction), d is an external disturbance and u is the control. The generalized force F is assumed to satisfy

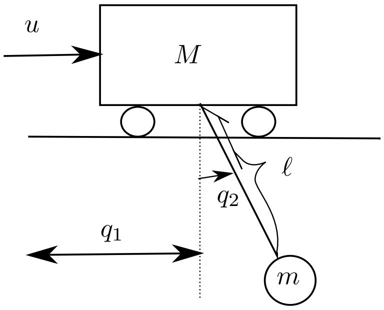
$$\dot{q}^T F(\dot{q}) \geq 0, \quad \dot{F}(0) = 0 \tag{10.5}$$

For control purposes (10.2) therefore takes the form

$$\frac{d}{dt}L_{\dot{q}}^T - L_q^T = -F(\dot{q}) + d + Bu \tag{10.6}$$

If the vectors u and q have the same size and B is nonsingular the system is said to be *fully actuated*, while the case of fewer components in u than q is called *under-actuated*.

Example 10.1 Consider a pendulum hanging from a cart.



The system is described by the position of the cart (q_1) and the angle of the pendulum (q_2). The kinetic energy is

$$T = \frac{1}{2}\begin{bmatrix}\dot{q}_1 & \dot{q}_2\end{bmatrix}\begin{bmatrix}M+m & m\ell\cos q_2 \\ m\ell\cos q_2 & m\ell^2\end{bmatrix}\begin{bmatrix}\dot{q}_1 \\ \dot{q}_2\end{bmatrix}$$

Here we have assumed the pendulum to be a point mass m at the end of the rod. The potential energy is given by

$$V = -mg\ell\cos q_2$$

If we assume that F has the form

$$F = \begin{bmatrix}b_1 & 0 \\ 0 & b_2\end{bmatrix}\dot{q}$$

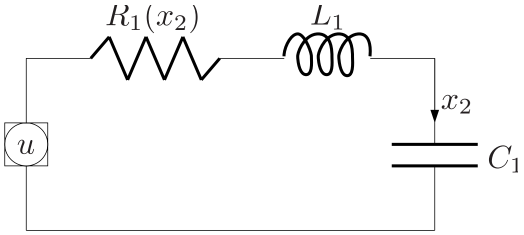
then the generalized force is

$$Q = -\begin{bmatrix}b_1 & 0 \\ 0 & b_2\end{bmatrix}\dot{q} + \begin{bmatrix}1 \\ 0\end{bmatrix}u$$

Note that this is an under-actuated system. ■

To show that the Lagrangian formalism can be used also for non-mechanical systems we consider an electrical circuit.

Example 10.2 A series circuuit is given by the following diagram



Let $q = x_1$ be the charge on the capacitor. Then $\dot{q} = x_2$ is the current into the capacitor which is also the common current all through the series circuit. Now T becomes the energy of the inductor

$$T = \frac{L_1}{2}\dot{q}^2$$

while V is the energy in the capacitor

$$V = \frac{1}{2C_1}q^2$$

The Lagrangian is then

$$L = \frac{L_1}{2}\dot{q}^2 - \frac{1}{2C_1}q^2$$

The generalized forces will correspond to voltages, composed of external signals and dissipation, i.e. voltage drops across resistors. In this case one gets

$$Q = u - R_1(\dot{q})$$

so that the dynamics is given by

$$\frac{d}{dt}(L_1\dot{q}) + \frac{1}{C_1}q = u - R_1(\dot{q})$$

and the state equations become

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{1}{L_1}(u - x_1/C_1 - R_1(x_2)) \end{aligned}$$

in accordance with circuit theory. ■

Equilibria

An equilibrium of a Lagrangian system is solution of (10.6) for $d = 0$, $u = 0$ where q is constant, and consequently $\dot{q}$ is identically zero. It follows that T is also identically zero so that (10.6) reduces to

$$V_q(q) = 0 \tag{10.7}$$

The equilibria are therefore the stationary points of the potential function.

Passivity and stability

Systems satisfying Lagrange's equations have several nice properties that are related to the following result.

Theorem 10.1 For a system satisfying

$$\frac{d}{dt}L_{\dot{q}}^T(q,\dot{q}) - L_q^T(q,\dot{q}) = -F(\dot{q}) + Bu \tag{10.8}$$

the following relation holds

$$\frac{d}{dt}H(q,\dot{q}) = -\dot{q}^T F(\dot{q}) + \dot{q}^T Bu \tag{10.9}$$

where $H(q,\dot{q}) = T(q,\dot{q}) + V(q)$ can be interpreted as the total energy.

Proof. Using the fact that $L_{\dot{q}}\dot{q} = 2T$ one has

$$\begin{aligned}\frac{d}{dt}H &= \frac{d}{dt}(T+V) = \frac{d}{dt}(2T - T + V) = \frac{d}{dt}(L_{\dot{q}}\dot{q} - L) = \\ &= \left(\frac{d}{dt}L_{\dot{q}}\right)\dot{q} + L_{\dot{q}}\ddot{q} - L_{\dot{q}}\ddot{q} - L_q\dot{q} = \left(\frac{d}{dt}L_{\dot{q}} - L_q\right)\dot{q} = \dot{q}^T(-F(\dot{q}) + Bu)\end{aligned}$$

■

This result has a number of consequences.

Corollary 10.1 Define the output to be $y = B^T\dot{q}$. Then the system described by (10.8) is passive.

Proof. Since V is assumed bounded from below, this is true also for H so that $H \geq c$ for some constant c. From (10.9) then follows that

$$\int_0^T y^T u dt + H(x(0)) = \dot{q}^T F(\dot{q}) + H(x(T)) \geq c$$

so that

$$\int_0^T y^T u dt + \gamma(x(0)) \geq 0$$

with $\gamma = H - c$. ■

Corollary 10.2 For a system (10.8) without external signals ($u = 0$) H is a Lyapunov function.

Proof. From (10.9) it follows that $\frac{d}{dt}H(q,\dot{q}) = -\dot{q}^T F(\dot{q}) \leq 0$ ■

Corollary 10.3 For a system (10.8) without external signals ($u = 0$) assume that the following holds, for some constants $\epsilon_1 > 0$, $\epsilon_2 > 0$.

- The potential V is radially unbounded.
- The kinteic energy satisfies $T \geq \epsilon_1\dot{q}^T\dot{q}$.
- For the potential V there is precisely one point q_o such that $V_q(q_o) = 0$
- The dissipative force satisfies $\dot{q}^T F(\dot{q}) \geq \epsilon_2\dot{q}^T\dot{q}$.

Then the equilibrioum $q = q_o, \dot{q} = 0$ is globally asymptotically stable.

Proof. From the first two conditions follows that H is radially unbounded. Using (10.9) gives

$$\frac{d}{dt}H(q,\dot{q}) = -\dot{q}^T F(\dot{q}) \leq 0$$

with equality only for $\dot{q} = 0$. If $\dot{q} = 0$ along a trajectory, then from (10.8) one must have $V_q(q) = 0$ which implies that $q = q_0$. Theorem 5.1 is then applicable.
■

Control of fully actuated Lagrangian systems

For a fully actuated system the matrix B is square and invertible. It is then no restriction to assume that B is a unit matrix so that the system dynamics is

$$\frac{d}{dt}L_{\dot{q}}^T(q,\dot{q}) - L_q^T(q,\dot{q}) = -F(\dot{q}) + u \tag{10.10}$$

Suppose one wants to control a Lagrangian system around a constant set point q_r. This means that q_r has to be an equilibrium point for the closed loop system. If one wants to keep the Lagrangian structure this means (according to (10.7)) that the potential has to be changed so that q_r becomes a stationary point. This can be done by defining a new potential W with a minimum at q_r and using the control

$$u = V_q^T(q) - W_q^T(q)$$

The system dynamics then becomes

$$\frac{d}{dt}\mathcal{L}_{\dot{q}}^T(q,\dot{q}) - \mathcal{L}_q^T(q,\dot{q}) = -F(\dot{q})$$

where $\mathcal{L} = T - W$. A simple choice of W is $W = \frac{1}{2}(q_r - q)^T K_p (q_r - q)$ for some positive definite matrix K_p. The control is then

$$u = V_q^T(q) + K_p(q_r - q)$$

To affect the convergence to the equilibrium the control can be extended to

$$u = V_q^T(q) + K_p(q_r - q) - K_d\dot{q} \tag{10.11}$$

where K_d is a positive definite matrix. The closed loop dynamics is then

$$\frac{d}{dt}\mathcal{L}_{\dot{q}}^T(q,\dot{q}) - \mathcal{L}_q^T(q,\dot{q}) = -(F(\dot{q}) + K_d\dot{q})$$

The effect of the $K_d\dot{q}$-term can thus be interpreted as an addition to the natural friction term F. Since K_D is positive definite the conditions of Corollary 10.3 will be met so that q_r is a globally asymptotically stable equilibrium. Note that the controller (10.11) that achieves this is a multivariable PD-controller with a feedforward from the potential energy term.

10.2 Interconnected systems

Many engineering systems consist of interconneted simple systems. This fact is reflected by modeling concepts like bond graphs and is used in object oriented modeling languages like Modelica. The Hamiltonian modeling techniques of classical physics can be adapted to cover systems of this type and also give a framework for design of controllers.

Storage elements

A basic idea is that energy is stored in simple elements that are linked together. Each storage element is assumed to have the following properties.

- There is a stored quantity x (e.g.electric charge in a capacitor).
- The time derivative of the stored quantity is regarded as a *flow* $f = \dot{x}$ into the component (e.g. electric current).
- The stored energy is a function $H(x)$ of the stored quantity (e.g. $\frac{1}{2C}x^2$ for a capacitor)
- There is defined an *effort* variable $e = \frac{dH}{dx}$ (e.g. the voltage of a capacitor)
- The power absorbed into the system is thus $\frac{d}{dt}H(x) = \frac{dH}{dx}\dot{x} = ef$

These properties are consistent with those of a C-element in bond graph theory. For an I-element the roles of effort and flow are reversed.

Now assume that there are n storage elements and that each element is described by a stored variable x_i, a flow f_i, an energy storage function $H_i(x_i)$ and en effort $e_i = dH_i/dx_i$. We introduce the vectors

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}, \quad e = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$$

For simplicity it is assumed that the total energy H is just the sum of the enrgies stored in the individual components (i.e. there are no phenomena like mutual incuctances between components).

$$H(x) = H_1(x_1) + \cdots + H_n(x_n) \tag{10.12}$$

The connection of the different components is assumed to give a linear relation

$$f = Me \tag{10.13}$$

for some matrix M. This is true in circuit theory (Kirchoff's laws) and for bond graphs (p- and s-junctions). It is assumed that the interconnection itself does not store, dissipate or generate energy. This means that the total power $e^T f$ going into the interconnection has to be zero so that, for all e,

$$0 = e^T f = e^T Me = e^T M^T e \Rightarrow e^T(M + M^T)e = 0 \Rightarrow M = -M^T$$

i. e. the matrix M has to be skew-symmetric.

The interconnected systems we have described are described by

$$\dot{x} = f = Me = MH_x(x)^T, \quad H_x = (\frac{\partial H}{\partial x_1}, \ldots, \frac{\partial H}{\partial x_n}) \tag{10.14}$$

with M skew-symmetric. Systems of this form are called *Hamiltonian* with *Hamilton function* H. Since $\dot{H} = H_x\dot{x} = H_xMH_x^T = 0$ the total energy is constant.

Now assume that some efforts and flows are not connected to storage elements but are inputs and outputs. Partition the vectors and M as

$$e = \begin{bmatrix} e_x \\ e_u \end{bmatrix}, \quad f = \begin{bmatrix} f_x \\ f_y \end{bmatrix}, \quad M = \begin{bmatrix} M_{xx} & M_{xu} \\ -M_{ux}^T & 0 \end{bmatrix}$$

where e_x, f_x are connected to storage elements, $e_u = u$ is the input and $f_y = -y$ is the output. The system description is then

$$\dot{x} = M_{xx}e_x + M_{xu}e_u = M_{xx}H_x^T(x) + M_{xu}u \tag{10.15}$$

$$y = -f_y = M_{ux}^T H_x^T(x) \tag{10.16}$$

10.3 Port controlled Hamiltonian systems

The systems described by (10.15) – (10.16) are a special case of so called *port controlled Hamiltonian systems*. In general they are systems that casn be written in the following form.

$$\begin{aligned} \dot{x} &= J(x)H_x^T(x) + g(x)u \\ y &= g^T(x)H_x^T(x) \end{aligned} \tag{10.17}$$

where $J(x)$ is skew-symmetric. This system satisfies

$$\frac{d}{dt}H(x) = H_x\dot{x} = H_xJ(x)H_x^T(x) + H_xg(x)u = y^Tu$$

showing that H is constant as long as $u = 0$. This means that systems having an internal dissipation of energy can not be modelled. One way to model dissipation is to consider a port controlled Hamiltonian system with two sets of inputs and outputs:

$$\begin{aligned} \dot{x} &= J(x)H_x^T(x) + g(x)u + g_R(x)u_R \\ y &= g^T(x)H_x^T(x) \\ y_R &= g_R^T(x)H_x^T(x) \end{aligned}$$

The input u_R and the output y_R are then connected by some mathematical relation $u_R = \phi(x, y_R)$. Assuming this relation to have the form $u_R = -\bar{R}(x)y_R$ with $\bar{R} \geq 0$, the model is then

$$\begin{aligned} \dot{x} &= (J(x) - R(x))H_x^T(x) + g(x)u \\ y &= g^T(x)H_x^T(x) \end{aligned} \tag{10.18}$$

where $R(x) = g(x)\bar{R}(x)g^T(x) \geq 0$. A model of the form (10.18), where J is a skew symmteric matrix and R a non-negative definite matrix, is called a *port controlled Hamiltonian system with dissipation.* The structure of the system immediately gives the following result.

Proposition 10.1 A system described by (10.18) is passive with the following energy balance.

$$\int_0^T y^T u\,dt + H(x(0)) - H(x(T)) = \int_0^T H_x R H_x^T\,dt \geq 0 \tag{10.19}$$

Proof. Follows from an integration of the equation

$$\frac{d}{dt}H = H_x\dot{x} = H_x J H_x^T - H_x R H_x^T + H_x g u = -H_x R H_x^T + y^T u$$

■

Example 10.3 Consider again the electrical system of Example 10.2. If x_1 denotes the charge on the capacitor, and x_2 the magnetic flow in the inductor, the energy functions are

$$H_1(x_1) = \frac{x_1^2}{2C_1}, \quad H_2(x_2) = \frac{x_2^2}{2L_1}$$

The flows are

$$f_1 = \dot{x}_1 = \text{current into capacitor}, \quad f_2 = \dot{x}_2 = \text{voltage over inductor}$$

while the efforts are

$$e_1 = \frac{dH_1}{dx_1} = \frac{x_1}{C_1}, \quad e_2 = \frac{dH_2}{dx_2} = \frac{x_2}{L_1}$$

If we assume that the voltage drop over the resistor has the form $r(e_2)e_2$, the relations between flows and efforts can be written

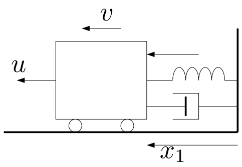
$$f_1 = e_2, \quad f_2 = -e_1 - r(e_2)e_2 + u$$

leading to the system equations

$$\dot{x} = \left(\underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_{J} - \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & r \end{bmatrix}}_{R} \right) \begin{bmatrix} H_{x_1} \\ H_{x_2} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

■

Example 10.4 Consider the mechanical system



where the spring is nonlinear with spring force $e_1 = x_1 + x_1^3$ and the damping is linear with force bv. If the state variable x_1 is the position, the spring energy is

$$H_1(x_1) = \frac{x_1^2}{2} + \frac{x_1^4}{4}$$

Let the mass be m and define $x_2 = mv$. Then the kinetic energy is

$$H_2(x_2) = \frac{x_2^2}{2m}, \quad e_2 = H_{x_2} = \frac{x_2}{m} = v$$

The flow variables are

$$f_1 = \dot{x}_1 = v, \quad f_2 = \dot{x}_2 = \text{total force}$$

with the relations

$$\begin{aligned} f_1 &= e_2 \\ f_2 &= u - e_1 - be_2 \end{aligned}$$

From these relations the following model can be deduced

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \left(\underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_{J(x)} - \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & b \end{bmatrix}}_{R(x)} \right) \underbrace{\begin{bmatrix} x_1 + x_1^3 \\ \frac{x_2}{m} \end{bmatrix}}_{H_x^T} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 + x_1^3 \\ \frac{x_2}{m} \end{bmatrix}$$

where $H = \frac{1}{2}x_1^2 + \frac{1}{4}x_1^4 + \frac{1}{2}x_2^2$ ■

Using the Hamiltonian structure for control

There are basically two things you can do, using state feedback, if you want to keep the Hamiltonian structure. The first one is to change the energy function H, the second one is to change the damping given by R. To change H one can use a control $u = k(x) + v$ satifying

$$(J(x) - R(x))\bar{H}_x^T = g(x)k(x) \tag{10.20}$$

The dynamics of (10.18) is then changed into

$$\dot{x} = (J(x) - R(x))(H + \bar{H})_x^T + g(x)v$$

where the Hamiltonian is changed from H to $H + \bar{H}$. To change R (or J) one can use a state feedback $u = k(x) + v$, where k satisfies

$$(\bar{J}(x) - \bar{R}(x))H_x^T = g(x)k(x) \tag{10.21}$$

The system dynamics is then changed into:

$$\dot{x} = (J(x) + \bar{J}(x) - R(x) - \bar{R}(x))H_x^T + g(x)v$$

The structures in equations (10.20) and (10.21) impose restrictions on what can be achieved by state feedback while keeping the Hamiltonian structure. This is seen in the following example.

Example 10.5 Consider again Example 10.4. Equation (10.20) becomes

$$\begin{aligned} \bar{H}_{x_2} &= 0 \\ -\bar{H}_{x_1} &= k(x) \end{aligned}$$

showing that $\bar{H}$ has to depend only on x_1. It is also clear from the second equation that the x_1-dependence of $\bar{H}$ can be chosen arbitrarily by choosing k suitably. It is therefore possible to place a stable equilibrium (a minimum of H) at any x_1-value. From (10.21) it is seen that $\bar{R}$ can be chosen to be of the form

$$\begin{bmatrix} 0 & 0 \\ 0 & \bar{b} \end{bmatrix}$$

giving the relation

$$k(x) = -\bar{b}\frac{x_2}{m}$$

The total control then has the form

$$k(x) = -\bar{H}_{x_1} - \bar{b}\frac{x_2}{m}$$

■