C21 Nonlinear Systems

4 Lectures Hilary Term 2023   Mark Cannon

1 Example class   mark.cannon@eng.ox.ac.uk

Contents

1 Introduction

1.1 Scope and objectives

The course aims to provide an overview of techniques for analysis and control design for nonlinear systems. Whereas linear system control theory is largely based on linear algebra (eg. when determining the behaviour of solutions to linear differential equations) and complex analysis (eg. when predicting system behaviour from transfer functions), the broader range of behaviour exhibited by nonlinear systems requires a wider variety of techniques. The course gives an introduction to some of the most useful and commonly used tools for determining system behaviour from a description in terms of differential equations. Two main topics are covered:

- **Lyapunov stability** — An intuitive approach to analyzing stability and convergence of dynamic systems without explicitly computing the solutions of their differential equations. This method forms the basis of much of modern nonlinear control theory and also provides a theoretical justification for using local linear control techniques.

- **Passivity and linearity** — These are properties of two important classes of dynamic system which can simplify the application of Lyapunov stability theory to interconnected systems. In this course we use the stability properties of linear and passive systems to derive the circle criterion, which provides an extension of the Nyquist criterion to nonlinear systems consisting of linear and nonlinear subsystems.

The theory discussed in this course is the basis of control design methods covered in C21, C20 and elsewhere. Some important topics *not* covered in this course are: input-output stability theory (which investigates stability using functional analysis) and differential geometry (which extends observability and controllability and provides the basis for feedback linearization).

Webpages: `weblearn.ox.ac.uk`

        `markcannon.github.io/teaching`

Please send corrections or suggestions to `mark.cannon@eng.ox.ac.uk`

At the end of the course you should be able to:

- Understand the basic Lyapunov stability definitions

- Analyse stability using the linearization method

- Analyse stability by Lyapunov's direct method

- Analyse convergence using Barbalat's Lemma

- Use invariant sets to determine regions of attraction

- Construct Lyapunov functions for linear and passive systems

- Use the circle criterion to analyse the stability of systems with static nonlinearities

1.2   Books

These notes are self-contained and they also contain some non-examinable background material (these sections are indicated in the text). For a full understanding of the course it will be helpful to read more detailed treatments given by textbooks on nonlinear control. Throughout these notes there are references to additional reading material.

1. J.-J. Slotine and W. Li, *Applied Nonlinear Control*, Prentice-Hall, 1991.

   Most of the material covered by the course is in chapters 3 and 4.

2. K.J. Astrom and R.M. Murray, *Feedback Systems: an introduction for scientists and engineers*, Princeton University Press, 2008.

   Chapter 4 introduces Lyapunov stability analysis.

3. M. Vidyasagar, *Nonlinear Systems Analysis (2nd ed.)*, Prentice-Hall, 1993.

   Chapter 5 discusses Lyapunov stability analysis. Chapter 2 has useful background material on nonlinear ODEs.

4. H.K. Khalil, *Nonlinear Systems (2nd ed.)*, Prentice-Hall, 1996.

   Chapters 1, 3, 4, 10 and 11 are relevant to this course.

1.3   Motivation [1]

There already exists a large amount of theory concerning control of linear systems, including robust and optimal control for multivariable linear systems of arbitrarily high order. So why bother designing controllers explicitly for nonlinear systems?

The main justification for nonlinear control is based on the observations:

- *All physical systems are nonlinear.* Apart from limitations on standard linear modeling assumptions such as Hooke's law, linearity of resistors and capacitors etc., nonlinearity invariably appears due to friction and heat dissipation effects.

- *Linearization is approximate.* Linear descriptions of physical processes are necessarily local (ie. accurate only within a restricted region of operation), and may be of limited use for control purposes.

Nonlinear control techniques are therefore useful when: (i) the required range of operation is large; (ii) the linearized model is inadequate (for example, linearizing the first order system $\dot{x} = xu$ about $x = 0$, $u = 0$ results in the uncontrollable system $\dot{x} = 0$). Further reasons for considering nonlinear controllers are: (iii) the ability of robust nonlinear controllers to tolerate large variations in uncertain system parameters; (iv) the simplicity of nonlinear control designs, which are often based on the physics of the process.

Linear vs. nonlinear system behaviour.   The analysis and control of nonlinear systems requires a different set of tools than can be used in the case of linear systems. In particular, for the linear system:

$$\dot{x} = Ax + Bu$$

where x is the state and u the control input, for the case of $u = 0$ we have:

- $x = 0$ is the unique equilibrium point (unless A is singular)

- stability is unaffected by initial conditions

and for $u \neq 0$:

---

[1]See *Slotine and Li* §1.2 pp4–12, for a more detailed discussion.

    – x remains bounded whenever u is bounded if $\dot{x} = Ax$ is stable

    – if $u(t)$ is sinusoidal, then in steady state $x(t)$ contains sinusoids of the same frequency as u

    – if $u = u_1 + u_2$, then $x = x_1 + x_2$.

However **none** of these properties is in general true for nonlinear systems.

Example 1.1. The system:

$$\ddot{x} + \dot{x} + k(x) = u, \qquad k(x) = x(x^2 + 1) \tag{1.1}$$

is a model of displacement $x(t)$ in a mass-spring-damper system subject to an externally applied force $u(t)$. The nonlinearity appears in the spring term $k(x)$, which stiffens with increasing $|x|$ (figure 1), and might therefore represent the effect of large elastic deflections.
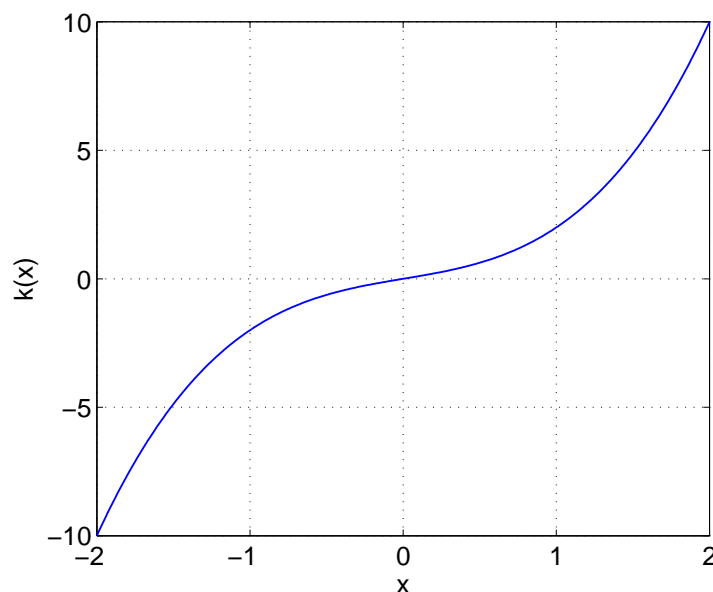

Figure 1: Spring characteristics $k(x)$ in (1.1).

Figure 2 shows the variation of x with time t in response to step changes in input u. Clearly the responses to positive steps are more oscillatory than those for negative step changes in u. This reduction in apparent damping ratio is due simply to the increase in stiffness of the nonlinear spring for large $|x|$. Note also that the steady state value of x for $u = 5$ is not 10 times that for $u = 50$, as would be expected in the linear case.   ◇
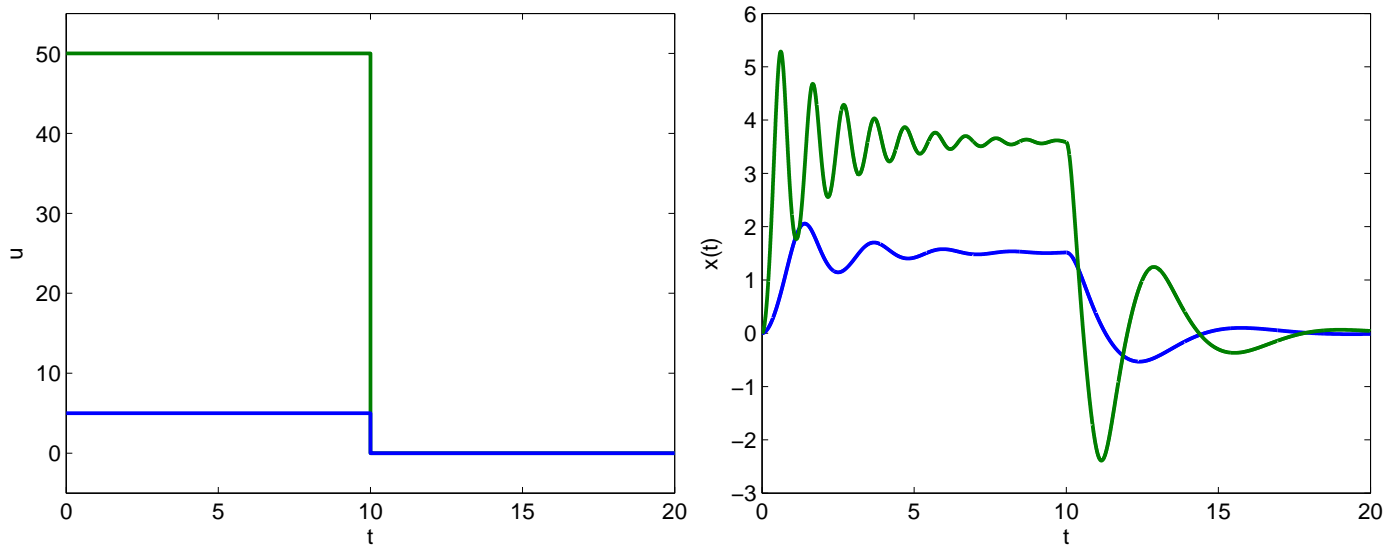
Figure 2: Response to steps in u of 5 (solid line), and 50 (dotted line).

Example 1.2. Van der Pol's equation:

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = 0, \qquad \mu > 0, \tag{1.2}$$

is a well-known example of a nonlinear system exhibiting limit cycle behaviour. In contrast to the limit cycles that occur in marginally stable linear systems, the amplitudes of limit cycle oscillations in nonlinear systems are independent of initial conditions. As a result the state trajectories of (1.2) all tend towards a single closed curve (the limit cycle), which can be seen in figure 3b.   ◊

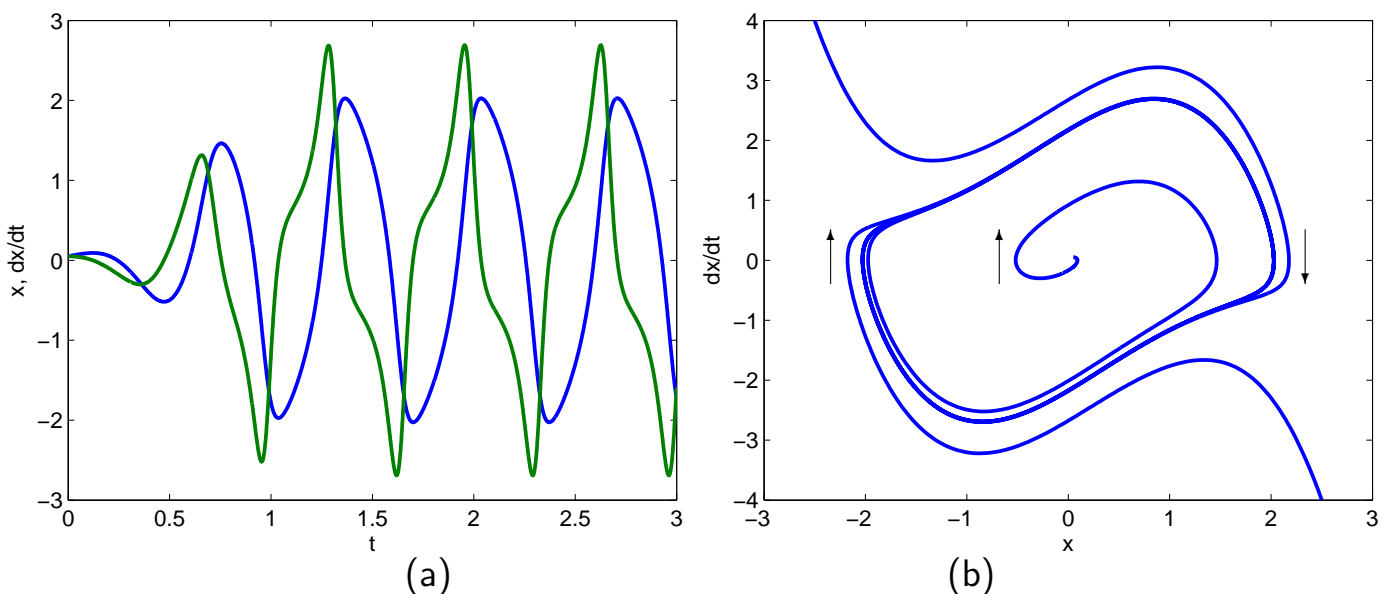(a)   (b)

Figure 3: Limit cycle of Van der Pol's equation (1.2) with $\mu = 0.5$. (a) Response for initial condition $x(0) = dx/dt(0) = 0.05$. (b) State trajectories.

2   Lyapunov stability

Throughout these notes we use state equations of the form

$$\dot{x} = f(x, u, t) \qquad \begin{cases} x : & \text{state variable} \\ u : & \text{control input} \end{cases} \qquad (2.1)$$

to represent systems of coupled ordinary differential equations. You should be familiar with the concept of state space from core course lectures, but to refresh your memory, suppose an nth-order system is given by

$$y^{(n)} = h\big(y, \dot{y}, \ldots y^{(n-1)}, u, t\big) \qquad (2.2)$$

(where $y^{(i)} = d^i y / dt^i$, $i = 1, 2, \ldots$), for some possibly nonlinear function h. Then an n-dimensional state vector x can be defined for example via

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} ; \qquad \begin{aligned} x_1 &= y \\ x_2 &= \dot{y} \\ &\vdots \\ x_{n-1} &= y^{(n-2)} \\ x_n &= y^{(n-1)} \end{aligned}$$

and (2.2) is equivalent to

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= h(x_1, x_2, \ldots x_n, u, t). \end{aligned}$$

This now has the form of (2.1), with

$$f(x, u, t) = \begin{bmatrix} x_2 \\ x_3 \\ \cdots \\ x_n \\ h(x_1, \ldots x_n, u, t) \end{bmatrix} .$$

We make a distinction between system dynamics which are invariant with time and those which depend explicitly on time.

> **Definition 2.1** (Autonomous/non-autonomous dynamics). The system (2.1) is **autonomous** if f does not depend explicitly on time t, ie. if (2.1) can be re-written $\dot{x} = f_1(x)$ for some function f_1. Otherwise the system (2.1) is **non-autonomous**.

For example, the closed-loop system formed by $\dot{x} = f(x, u)$ under time-invariant feedback control $u = u(x)$ is given by $\dot{x} = f(x, u(x))$, and is therefore autonomous. However a system $\dot{x} = f(x, u)$ under time-varying feedback $u = u(x, t)$ (which would be needed if $x(t)$ were required to follow a time-varying target trajectory) has closed-loop dynamics $\dot{x} = f(x, u(x, t))$ which are therefore non-autonomous.

As stated earlier, a nonlinear system may have different stability properties for different initial conditions. It is therefore usual to consider stability and convergence with respect to an equilibrium point, defined as follows.

> **Definition 2.2** (Equilibrium). A state x^* is an equilibrium point if $x(t_0) = x^*$ implies that $x(t) = x^*$ for all $t \geq t_0$.

Clearly the equilibrium points of the forced system (2.1) are dependent on the control input $u(t)$. In fact the problem of controlling (2.1) so that $x(t)$ converges to a given target state x_d is equivalent to finding a control law $u(t)$ which forces x_d to be a stable equilibrium (in some sense) of the closed-loop system. Thus it is convenient consider the equilibrium points of an **unforced** system:

$$\dot{x} = f(x, t), \tag{2.3}$$

(which could of course be obtained from a forced system under a specified control input $u(t)$). By definition, an equilibrium point x^* of (2.3) is a solution of

$$f(x^*, t) = 0.$$

Note that solving for x^* may not be trivial for general f. The remainder of section 2 considers the stability of equilibrium points of (2.3). Following the

usual convention, we define f in (2.3) so that the origin $x = 0$ is an equilibrium point, ie. so that $f(0, t) = 0$ for all t.[2]

The origins of modern stability theory date back to Lagrange (1788), who showed that, in the absence of external forces, an equilibrium of a conservative mechanical system is stable if it corresponds to a local minimum of the potential energy stored in the system. Stability theory remained restricted to conservative dynamics described by Lagrangian equations of motion until 1892, when the Russian mathematician A. M. Lyapunov developed methods applicable to arbitrary differential equations. Lyapunov's work was largely unknown outside Russia until about 1960, when it received widespread attention through the work of La Salle and Lefschetz. With several refinements and modifications, Lyapunov's methods have become indispensable tools in nonlinear system control theory.

2.1   Stability definitions [3]

As might be expected, the most basic form of stability is simply a guarantee that the state trajectories starting from points in the vicinity of an equilibrium point remain close to that equilibrium point at all future times. In addition to this we consider the stronger properties of asymptotic and exponential stability, which ensure convergence of trajectories to an equilibrium point. Although the definitions discussed in this section are mostly intuitively obvious, they are often useful in their own right, particularly in cases where the stability theorems described in section 2.3 cannot be applied directly.

> **Definition 2.3** (Stability)**.** The equilibrium $x = 0$ of (2.3) is **stable** if, for each time t_0, and for every constant $R > 0$, there exists some $r(R, t_0) > 0$ such that
> $$\|x(t_0)\| < r \implies \|x(t)\| < R, \ \forall t \geq t_0.$$
> (Here $\|\cdot\|$ can be any vector norm.) It is **uniformly stable** if r is independent of t_0. The equilibrium is **unstable** if it is not stable.

---

[2]Any equilibrium x^* can be translated to the origin by redefining the state x as $x' = x - x^*$.
[3]See *Slotine and Li* §3.2 pp47–52 and §4.1 pp101–105, or *Vidyasagar* §5.1 pp135–147.

An equilibrium is therefore stable if $x(t)$ can be contained within an arbitrarily small region of state space for all $t \geq t_0$ provided $x(t_0)$ is sufficiently close to the equilibrium point (see figure 4).
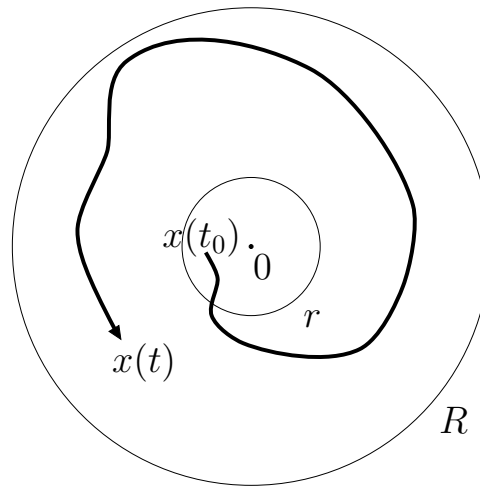

Figure 4: Stable equilibrium.

Note that:

1. Equilibrium points of stable autonomous systems are necessarily uniformly stable. This is because the state trajectories, $x(t)$, $t \geq t_0$, of an autonomous system depend only on initial condition $x(t_0)$ and not on initial time t_0.

2. An equilibrium point $x = 0$ may be unstable even though trajectories starting from points close to $x = 0$ do not tend to infinity. This is the case for Van der Pol's equation (example 1.2), which has an unstable equilibrium at the origin. (All trajectories starting from points within the limit cycle of figure 3b eventually join the limit cycle, and therefore it is not possible to find $r > 0$ in definition 2.3 whenever R is small enough that some points on the closed curve of the limit cycle lie outside the set of points x satisfying $\|x\| < R$.)

It is also worth noting that stability (as opposed to uniform stability) is a very weak condition which implies that an equilibrium point actually tends towards instability as $t \to \infty$. This is because it is only necessary to specify r in definition 2.3 as a function of t_0 if, for fixed R, $r(R, t_0)$ tends to zero as a function of t_0. Otherwise $r(R)$ could be specified independently of t_0 as the minimum value of $r(R, t_0)$ over all t_0.

Definition 2.4 (Asymptotic stability)**.** The equilibrium $x = 0$ of (2.3) is **asymptotically stable** if: (a) it is stable, and (b) for each time t_0 there exists some $r(t_0) > 0$ such that

$$\|x(t_0)\| < r \implies \|x(t)\| \to 0 \text{ as } t \to \infty.$$

It is **uniformly asymptotically stable** if it is asymptotically stable and both r and the rate of convergence in (b) are independent of t_0.

Asymptotic stability therefore implies that the trajectories starting from any point within some region of state space containing the equilibrium point remain bounded and converge asymptotically to the equilibrium (see figure 5).
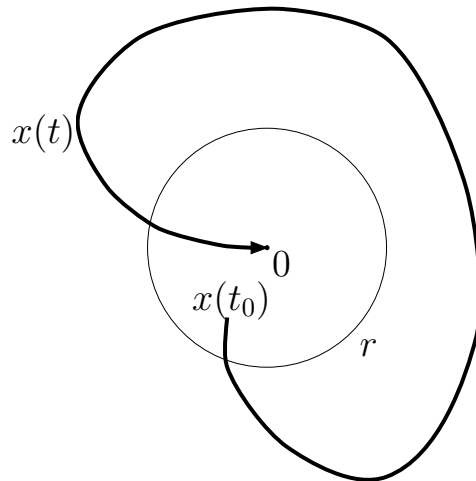
Figure 5: Asymptotically stable equilibrium.

Note that:

1. An asymptotically stable equilibrium of an autonomous system is necessarily uniformly asymptotically stable.

2. The convergence condition (b) of definition 2.4 is equivalent to requiring that, for every constant $R > 0$ there exists a $T(R, r, t_0)$ such that

$$\|x(t_0)\| < r \implies \|x(t)\| < R, \ \forall t \geq t_0 + T,$$

and the convergence rate is independent of t_0 if T is independent of t_0.

Example 2.5. The first order system

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

has general solution

$$x(t) = \frac{1+t_0}{1+t}x(t_0), \quad t \geq t_0.$$

Here $x = 0$ is uniformly stable (check the condition of definition 2.3 for the choice $r = R$), and all trajectories converge asymptotically to 0; therefore the origin is an asymptotically stable equilibrium. However the rate of convergence is dependent on initial time t_0 — if $\|x(t_0)\| < r$ then the condition $\|x(t)\| < R, \ \forall t \geq t_0 + T$ requires that $T > (1 + t_0)(r - R)/R$, which, for fixed R, cannot be bounded by any finite constant for all $t_0 \geq 0$. Hence the origin is not uniformly asymptotically stable.   ◇

Definition 2.6 (Exponential stability). The equilibrium $x = 0$ of (2.3) is exponentially stable if there exist constants $r, R, \alpha > 0$ such that

$$\|x(t_0)\| < r \implies \|x(t)\| \leq Re^{-\alpha t}, \ \forall t \geq t_0.$$

Note that:

1. Asymptotic and exponential stability are **local** properties of a dynamic system since they only require that the state converges to zero from a finite set of initial conditions (known as a **region of attraction**): x where $\|x\| < r$.

2. If r can be taken to be infinite in definition 2.4 or definition 2.6, then the system is respectively **globally asymptotically stable** or **globally exponentially stable**.[4]

3. A strictly stable linear system is necessarily globally exponentially stable.

---

[4]There is no ambiguity in talking about global stability of the overall system rather than global stability of a particular equilibrium point since a globally asymptotically or globally exponentially stable system can only have a single equilibrium point.

2.2   Lyapunov's linearization method [5]

In many cases it is possible to determine whether an equilibrium of a nonlinear system is locally stable simply by examining the stability of the linear approximation to the nonlinear dynamics about the equilibrium point. This approach is known as Lyapunov's linearization method since its proof is based on the more general stability theory of Lyapunov's direct method. However the idea behind the approach is intuitively obvious: within a region of state space close to the equilibrium point, the difference between the behaviour of the nonlinear system and that of its linearized dynamics is small since the error in the linear approximation is small for states close to the equilibrium.

The **linearization** of a system

$$\dot{x} = f(x), \qquad f(0) = 0, \tag{2.4}$$

about the equilibrium $x = 0$ is derived from the Taylor's series expansion of f about $x = 0$. Provided $f(x)$ is continuously differentiable[6] we have

$$\dot{x} = Ax + \overline{f}(x),$$

where A is the Jacobian matrix of f evaluated at $x = 0$:

$$A = \left[\frac{\partial f}{\partial x}\right]_{x=0} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}_{x=0}, \tag{2.5}$$

and by Taylor's theorem the sum of higher order terms, $\overline{f}$, satisfies

$$\lim_{x \to 0} \frac{\|\overline{f}(x)\|_2}{\|x\|_2} = 0, \ \forall t \geq 0. \tag{2.6}$$

Neglecting higher order terms gives the linearization of (2.4) about $x = 0$:

$$\dot{x} = Ax. \tag{2.7}$$

---

[5]See *Slotine and Li* §3.3 pp53–57 or *Vidyasagar* §5.5 pp209–219.

[6]A function is continuously differentiable if it is continuous and has continuous first derivatives.

The stability of the linear approximation can easily be determined from the eigenvalues of A. This is because the general solution of a linear system can be computed explicitly:

$$\dot{x} = Ax \implies x(t) = e^{A(t-t_0)}x(t_0),$$

and it follows that the linear system is strictly stable if and only if all eigenvalues of A have negative real parts (or marginally stable if the real parts of some eigenvalues of A are equal to zero, and the rest are negative).

> **Theorem 2.7** (Lyapunov's linearization method)**.** *For the nonlinear system (2.4), suppose that f is continuously differentiable and define A as in (2.5). Then:*
> - *$x = 0$ is an **exponentially stable** equilibrium of (2.4) if all eigenvalues of A have negative real parts.*
> - *$x = 0$ is an **unstable** equilibrium of (2.4) if A has at least one eigenvalue with positive real part.*

The proof of the theorem makes use of (2.6), which implies that the error $\overline{f}$ in the linear approximation to f converges to zero faster than any linear function of x as x tends to zero. Consequently the stability (or instability) of the linearized dynamics implies local stability (or instability) of the equilibrium point of the original nonlinear dynamics.

Note that:

1. The linearization approach concerns local (rather than global) stability.

2. If the linearized dynamics are marginally stable then the equilibrium of the original nonlinear system could be either stable or unstable (see example 2.8 below). It is not possible to draw any conclusions about the stability of the nonlinear system from the linearization in this case since the local stability of the equilibrium could be determined by higher order terms that are neglected in the linear approximation.

3. The above analysis can be extended to non-autonomous systems of the form (2.3). However the linearized system is then time-varying (ie. of

the form $\dot{x} = A(t)x$), and its stability is therefore more difficult to determine in general. In this case the equilibrium of the nonlinear system is asymptotically stable if the linearized dynamics are asymptotically stable.

Example 2.8. Consider the following first order system

$$\dot{x} = -\alpha x |x|$$

where α is a constant. For $\alpha > 0$, the derivative $\dot{x}(t)$ is of opposite sign to $x(t)$ at all times t, and the system is therefore globally asymptotically stable. If $\alpha < 0$ on the other hand, then $\dot{x}(t)$ has the same sign as $x(t)$ for all t, and in this case the system is unstable (in fact the general solution shows that $|x(t)| \to \infty$ as $t \to \alpha^{-1}(t_0 + 1/|x(t_0)|)$). However the linearized system is given by

$$\dot{x} = 0$$

which is marginally stable irrespective of the value of α, and therefore contains no information on the stability of the nonlinear system. $\Diamond$

The linearization of a forced system $\dot{x} = f(x, u)$ under a given feedback control law $u = u(x)$ is most easily computed by directly neglecting higher order terms in the Taylor's series expansions of f and u about $x = 0, u = 0$. Thus

$$\dot{x} \approx Ax + Bu, \qquad A = \left[\frac{\partial f}{\partial x}\right]_{x,u=0} \quad B = \left[\frac{\partial f}{\partial u}\right]_{x,u=0} \tag{2.8a}$$

$$u \approx Fx, \qquad F = \left[\frac{\partial u}{\partial x}\right]_{x,u=0} \tag{2.8b}$$

and the linearized dynamics are therefore

$$\dot{x} = (A + BF)x. \tag{2.9}$$

The implication of Lyapunov's linearization method is that linear control techniques can be used to design locally stabilizing control laws for nonlinear systems via linearization about the equilibrium of interest. All that is required is a linear control law $u = Fx$, computed using the linearized dynamics (2.8a), which forces the linearized closed-loop system (2.9) to be strictly stable. However this approach is limited by the local nature of the linearization process —

there is no guarantee that the resulting linear control law stabilizes the non-linear system everywhere within the desired operating region of state space, which may be much larger than the region on which the linear approximation to the nonlinear system dynamics is accurate.

2.3 Lyapunov's direct method [7] [8]

The aim of Lyapunov's direct method is to determine the stability properties of an equilibrium point of an unforced nonlinear system without solving the differential equations describing the system. The basic approach involves constructing a scalar function V of the system state x, and considering the derivative $\dot{V}$ of V with respect to time. If V is positive everywhere except at the equilibrium $x = 0$, and if furthermore $\dot{V} \leq 0$ for all x (so that V cannot increase along the system trajectories), then it is possible to show that $x = 0$ is a stable equilibrium. This is the main result of Lyapunov's direct method, and a function V with these properties is known as a **Lyapunov function**. By imposing additional conditions on V and its derivative $\dot{V}$, this result can be extended to provide criteria for determining whether an equilibrium is asymptotically or exponentially stable both locally and globally.

Motivation

The function V can be thought of as a generalization of the idea of the stored energy in a system. Consider for example the second order system:

$$m\ddot{y} + c(\dot{y}) + k(y) = 0 \qquad \begin{cases} m > 0 \\ c(0) = 0, \ \text{sign}\big(c(\dot{y})\big) = \text{sign}\,(\dot{y}) \\ k(0) = 0, \ \text{sign}\big(k(y)\big) = \text{sign}\,(y) \end{cases} \qquad (2.10)$$

where y is the displacement of a mass m, and $c(\dot{y})$, $k(y)$ are respectively nonlinear damping and spring forces acting on the mass (figure 6).

---

[7]See *Slotine and Li* §3.4 pp57–68 and §4.2 pp105–113 or *Vidyasagar* §5.3 pp157–176.

[8]The proofs of the various results given in this section are intended to be conceptual rather than rigorous. For a more technical treatment see eg. *Vidyasagar*, though this is not necessary for the level of understanding required in this course.
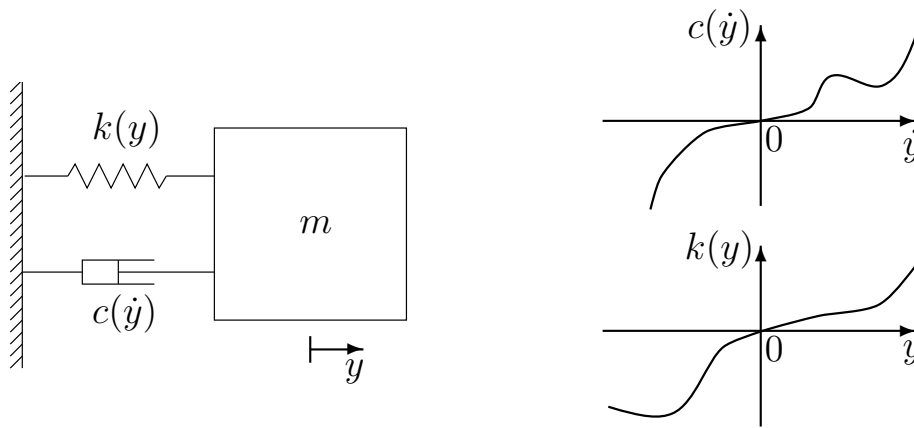
Figure 6: Mass, spring and damper.

From the rate of change of stored energy it is easy to deduce that the equilibrium $y = \dot{y} = 0$ is stable without explicit knowledge of the functions k and c. When released from non-zero initial conditions energy is transfered between the spring and the mass (since the spring force opposes the displacement of the mass), but the total stored energy decreases monotonically over time due to dissipation in the damper (since the damping force opposes the velocity of the mass). It is also intuitively obvious that a small value for the stored energy at time t corresponds to small values of $y(t)$ and $\dot{y}(t)$, and therefore $y = \dot{y} = 0$ must be a stable equilibrium.

To make this argument more precise, let V be the energy stored in the system (ie. the sum of the kinetic energy of the mass and the potential energy of the spring):

$$V(y, \dot{y}) = \frac{1}{2}m\dot{y}^2 + \int_0^y k(y)\,dy. \tag{2.11}$$

Here $k(y)$ has the same sign as y, and V is therefore strictly positive unless $y = \dot{y} = 0$ (in which case $V = 0$). It follows that $(y^2 + \dot{y}^2)^{1/2}$ can only exceed a given bound R if V increases beyond some corresponding bound $\overline{V}$. However, using the chain rule and the system dynamics (2.10), the derivative of V with respect to time is given by

$$\dot{V}(y, \dot{y}) = \frac{1}{2}m\frac{d}{d\dot{y}}\left[\dot{y}^2\right]\ddot{y} + \frac{d}{dy}\left[\int_0^y k(y)\,dy\right]\dot{y}$$
$$= m\dot{y}\ddot{y} + k(y)\dot{y}$$
$$= -c(\dot{y})\dot{y}$$

which implies that $\dot{V} \leq 0$ due to the condition on the sign of $c(\dot{y})$ in (2.10). Hence V decreases monotonically along system trajectories. It is therefore clear that if the initial conditions $y(t_0), \dot{y}(t_0)$ are sufficiently close to the equilibrium that $\left(y^2(t_0) + \dot{y}^2(t_0)\right)^{1/2} < r$, where r satisfies

$$(y^2 + \dot{y}^2)^{1/2} < r \implies V(y, \dot{y}) < \overline{V},$$

then $(y^2 + \dot{y}^2)^{1/2}$ cannot exceed the bound R at any future time $t \geq t_0$, which implies that the equilibrium is uniformly stable in the sense of definition 2.3. Figure 7 illustrates the relationship between the constants $\overline{V}$, r and R. To simplify the figure we have assumed that V has radial symmetry (ie. $V(y, \dot{y}) = V(y^2 + \dot{y}^2)$).
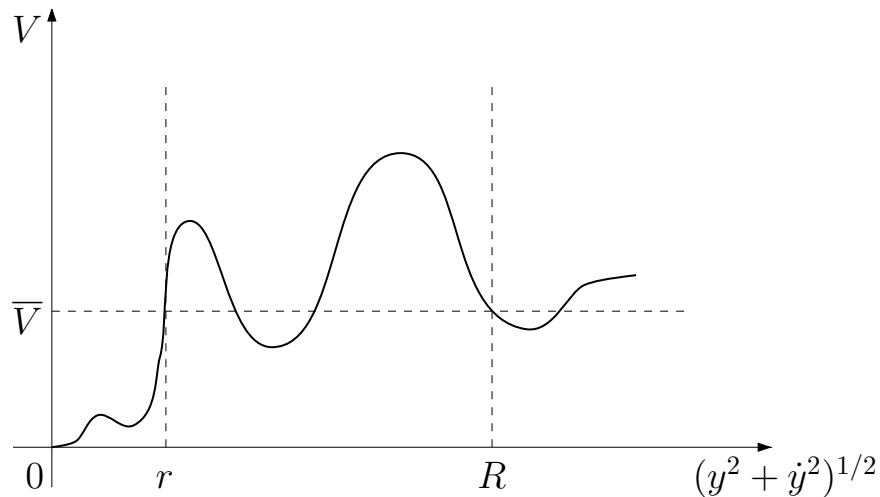

Figure 7: Storage function V and parameters $\overline{V}$, r and R.

Autonomous systems

Positive definite functions. The argument used above to show that the equilibrium of (2.10) is stable relies on the fact that a bound on V implies a corresponding bound on the norm of the state vector and vice-versa. This is the case for a general continuous function V of the state x provided V has the property that

$$x \neq 0 \iff V(x) > 0,$$
$$x = 0 \iff V(x) = 0.$$
$$\tag{2.12}$$

A continuous function with this property is said to be **positive definite**.[9]
Alternatively, if $V(x)$ is continuous and (2.12) holds for all x such that $\|x\| < R_0$ for some $R_0 > 0$, then V is **locally positive definite** (or positive definite for $\|x\| < R_0$).

Derivative of $V(x)$ along system trajectories. If x satisfies the differential equation $\dot{x} = f(x)$, then the time-derivative of a continuously differentiable scalar function $V(x)$ is given by

$$\dot{V}(x) = \nabla V(x)\dot{x} = \nabla V(x)f(x), \tag{2.13}$$

where $\nabla V(x)$ is the gradient of V (expressed as a row vector) with respect to x evaluated at x. The expression (2.13) gives the rate of change of V as x moves along a trajectory of the system state, and $\dot{V}$ is therefore known as the derivative of V along system trajectories.

With these definitions we can give the general statement of Lyapunov's direct method for autonomous systems.

Theorem 2.9 (Stability/asymptotic stability for autonomous systems). *If there exists a continuously differentiable scalar function $V(x)$ such that:*

    (a). $V(x)$ is positive definite
    (b). $\dot{V}(x) \leq 0$

*for all x satisfying $\|x\| < R_0$ for some constant $R_0 > 0$, then the equilibrium $x = 0$ is **stable**. If, in addition,*

    (c). $-\dot{V}(x)$ is positive definite

*whenever $\|x\| < R_0$, then $x = 0$ is **asymptotically stable**.*

The first part of this theorem can be proved by showing that it is always possible to find a positive scalar r which ensures that, for any given $R > 0$, $x(t)$ is bounded by $\|x(t)\| < R$ for all $t \geq t_0$ whenever the initial condition satisfies $\|x(t_0)\| < r$. To do this, first choose $R < R_0$ and define $\overline{V}$ as the minimum value of $V(x)$ over all x such that $\|x\| = R$ (figure 8a). Then a trajectory $x(t)$ can escape the region on which $\|x\| < R$ only if $V(x(t)) \geq \overline{V}$

---

[9]Similarly, V is negative definite if $-V$ is positive definite.

for some $t \geq t_0$. But condition (a) implies that there exists a positive $r < R$ for which $V(x) < \overline{V}$ whenever $\|x\| < r$, whereas condition (b) ensures that $V\big(x(t)\big)$ decreases over time if $\|x(t)\| < R$. It follows that $V\big(x(t)\big)$ cannot exceed $\overline{V}$ for all $t \geq t_0$ if the initial state satisfies $\|x(t_0)\| < r$, and the state $x(t)$ cannot therefore leave the region on which $\|x\| < R$. For the case of $R \geq R_0$, simply repeat this argument with R replaced by R_0.
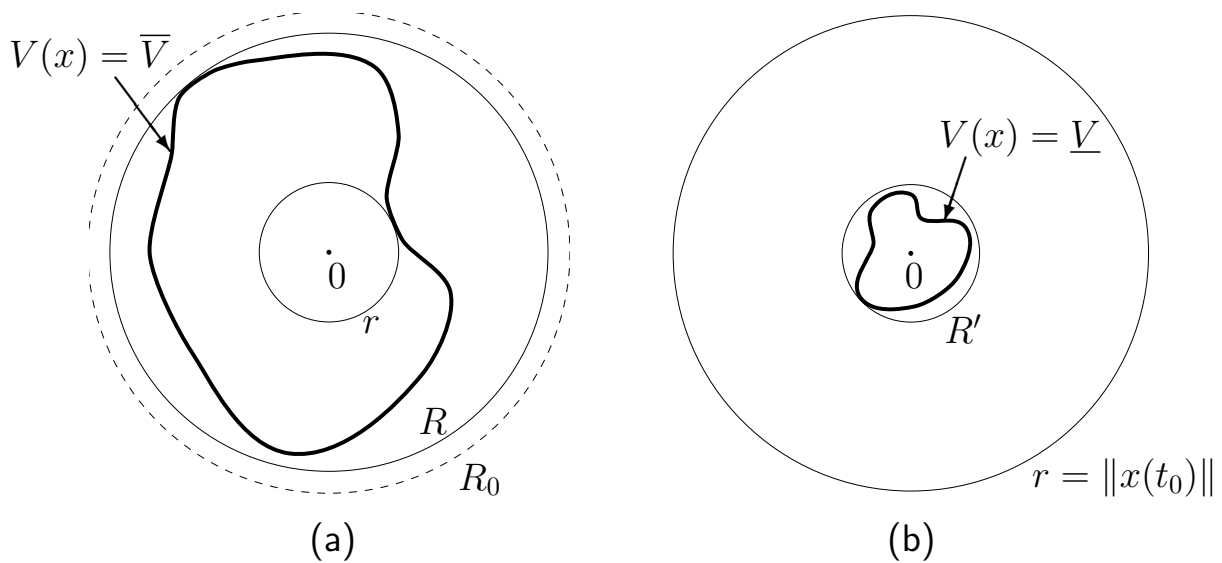

Figure 8: Level sets of $V(x)$ in state space.

The second part concerning asymptotic stability can be proved by contradiction. Suppose that the origin is not asymptotically stable, then for any given initial condition $x(t_0)$ the value of $\|x(t)\|$ must remain larger than some positive number $R' < \|x(t_0)\|$. Conditions (a) and (c) therefore imply that $V\big(x(t)\big) > \underline{V}$ and $\dot{V}\big(x(t)\big) < -\underline{W}$ at all times $t \geq t_0$, for some constants $\underline{V}, \underline{W} > 0$ (figure 8b). But this is a contradiction since $V\big(x(t)\big)$ must decrease to a value less than $\underline{V}$ in a finite time smaller than $[V\big(x(t_0)\big) - \underline{V}]/\underline{W}$ if $\dot{V}\big(x(t)\big) < -\underline{W}$ for all $t \geq t_0$.

Non-autonomous systems [10]

The discussion so far has concerned only autonomous systems (systems with dynamics of the form $\dot{x} = f(x)$). Extensions to non-autonomous systems ($\dot{x} = f(x,t)$) are straightforward, but the situation is complicated by the fact that a Lyapunov function for non-autonomous dynamics may need to be time-varying, ie. of the form $V(x,t)$. For completeness this section gives some results on non-autonomous system stability.

Positive definite time-varying functions. To extend the definition of a positive definite function to the case of time-varying functions, we simply state that $V(x,t)$ is positive definite if there exists a time-invariant positive definite function $V_0(x)$ satisfying

$$V(x,t) \geq V_0(x), \ \forall t \geq t_0, \ \forall x. \tag{2.14}$$

Similarly, V is locally positive definite if $V(x,t)$ is bounded below by a locally positive definite time-invariant function $V_0(x)$ for all $t \geq t_0$.

Decrescent functions. Non-autonomous systems differ from autonomous systems in that the stability of an equilibrium may be uniform or non-uniform (recall that the stability properties of autonomous systems are necessarily uniform). In order to ensure uniform stability, a Lyapunov function V for a non-autonomous system must also be **decrescent**, which requires that

$$V(x,t) \leq V_0(x), \ \forall t \geq t_0, \ \forall x \tag{2.15}$$

for some time-invariant positive definite function $V_0(x)$.

Derivative of $V(x,t)$ along system trajectories. If $x(t)$ satisfies $\dot{x} = f(x,t)$, then the derivative with respect to t of a continuously differentiable function $V(x,t)$ can be expressed

$$\dot{V}(x,t) = \frac{\partial V}{\partial t}(x,t) + \nabla V(x)\dot{x} = \frac{\partial V}{\partial t}(x,t) + \nabla V(x)f(x,t). \tag{2.16}$$

The appearance here of the partial derivative of V with respect to t is due to the explicit dependence of V on time.

---

[10]The discussion in this section of Lyapunov's direct method for non-autonomous systems is non-examinable and is provided as background material

The main result of Lyapunov's direct method for non-autonomous systems can be stated in terms of these definitions as follows.

Theorem 2.10 (Stability/asymptotic stability for non-autonomous systems)**.**
If there exists a continuously differentiable scalar function $V(x,t)$ such that:

    (a). $V(x,t)$ is positive definite

    (b). $\dot{V}(x,t) \leq 0$

for all x satisfying $\|x\| < R_0$ for some constant $R_0 > 0$, then the equilibrium $x = 0$ is **stable**. *If, furthermore,*

    (c). $V(x,t)$ is decrescent

then $x = 0$ is **uniformly stable**. *If, in addition to (a), (b), and (c),*

    (d). $-\dot{V}(x,t)$ is positive definite

whenever $\|x\| < R_0$, then $x = 0$ is **uniformly asymptotically stable**.

This theorem can be proved in a similar way to theorem 2.9. The only difference in the first part concerning stability is that here the definitions of the constants $\overline{V}$ and r must hold for all $t \geq t_0$ (ie. so that $\overline{V}$ is the minimum of $V(x,t)$ for all x such that $\|x\| = R$ and all $t \geq t_0$, and likewise $V(x,t) < \overline{V}$ for all x such that $\|x\| < r$ and all $t \geq t_0$). Similarly, in the second part concerning asymptotic stability, the constants $\underline{V}$ and $\underline{W}$ must be defined as lower bounds on $V(x,t)$ and $-\dot{V}(x,t)$ for all x such that $\|x\| < R'$ and all $t \geq t_0$. In order to show uniform stability, the requirement that V is decrescent in (c) avoids the possibility that $\overline{V}$ tends to infinity as $t \to \infty$ (which would imply that r becomes arbitrarily small as $t \to \infty$) by preventing $V(x,t)$ from becoming infinite within the region on which $\|x\| < R$ at any time $t \geq t_0$. Uniform asymptotic stability is implied by $V\big(x(t_0),t_0\big)$ necessarily being finite (so that the time taken for the value of V to fall below $\underline{V}$ is finite), due to the assumption that V is decrescent.

Global Stability

Lyapunov's direct method also provides a means of determining whether a system is globally asymptotically stable via simple extensions of the theorems for asymptotic stability already discussed. Before giving the details of this

approach, we first need to introduce the concept of a radially unbounded function.

Radially unbounded functions. As might be expected, the conditions of theorems 2.9 or 2.10 must hold at all points in state space in order to assert global asymptotic stability for autonomous and non-autonomous systems respectively. However one extra condition on V, which for time-invariant functions can be expressed:

$$V(x) \to \infty \text{ as } \|x\| \to \infty, \tag{2.17}$$

is also required. A function with this property is said to be **radially unbounded**. For time-varying functions this condition becomes

$$V(x,t) \geq V_0(x), \ \forall t \geq t_0, \ \forall x, \tag{2.18}$$

where V_0 is any radially unbounded time-invariant function.

Below we give the global version of theorems 2.9 and 2.10.

> **Theorem 2.11** (Global uniform asymptotic stability). *If V is a Lyapunov function for an autonomous system (or non-autonomous system) which satisfies conditions (a)–(c) of theorem 2.9 (or conditions (a)–(d) of theorem 2.10 respectively) for all x (ie. in the limit as $R_0 \to \infty$), and V is radially unbounded, then the system is **globally uniformly asymptotically stable**.*

The conditions on V in this theorem ensure that the origin is uniformly asymptotically stable, by theorem 2.9 (or theorem 2.10). To prove the theorem it therefore suffices to show that the conditions on V imply that every state trajectory $x(t)$ tends to the origin as $t \to \infty$. But since $x(t)$ is continuous, this requires that $x(t)$ remains bounded for all $t \geq t_0$ for arbitrary initial conditions $x(t_0)$. Hence the purpose of the radial unboundedness condition on V is to ensure that $x(t)$ remains at all times within the *bounded* region defined by $V(x,t) \leq V(x(t_0),t_0)$. If V were not radially unbounded, then not all contours of constant V in state-space would be closed curves, and it would be possible for $x(t)$ to drift away from the equilibrium even though $\dot{V}$ is negative.

The remainder of the proof involves constructing a finite bound on the time taken for a trajectory starting from arbitrary $x(t_0)$ to enter the region on which $\|x\| < R$, for any $R > 0$. This can be done by finding bounds on V and $-\dot{V}$ using the positive definite properties of V and $-\dot{V}$.

Note that:

1. If the requirement that $\dot{V}$ is negative definite in theorem 2.11 is replaced by the condition that $\dot{V}$ is simply non-positive, then $x(t)$ is guaranteed to be **globally bounded**. This means that, for every initial condition $x(t_0)$, there exists a finite constant $R\big(x(t_0)\big)$ such that $\|x(t)\| < R$ for all $t \geq t_0$.
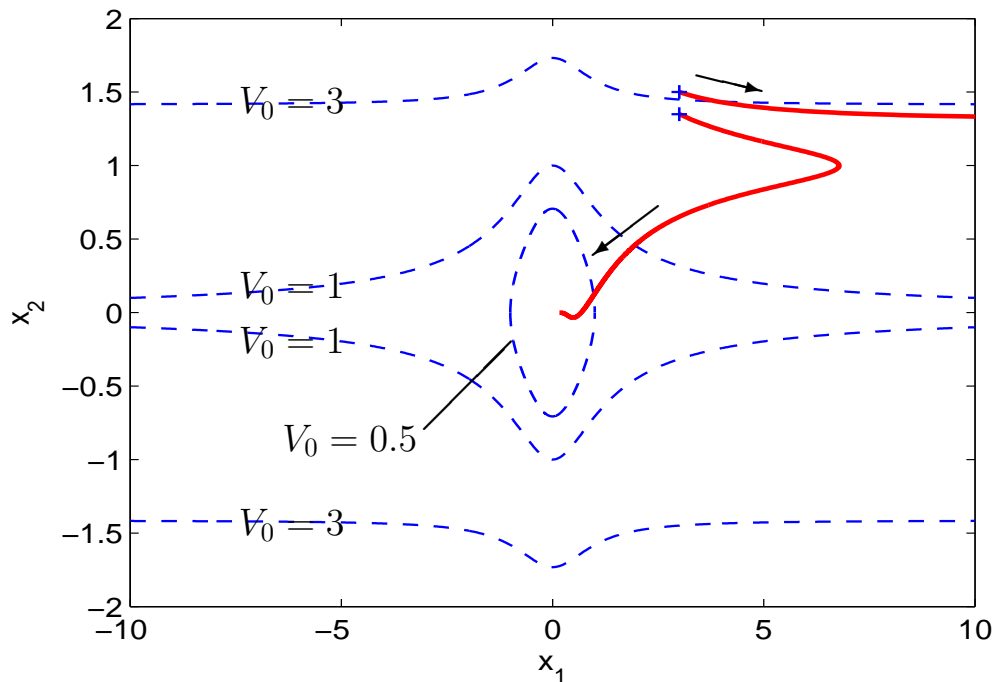

Figure 9: Contours of $V(x_1, x_2) = V_0$ for the function V defined in (2.20) (dashed lines), and state trajectories of (2.19) for two different initial conditions (solid lines).

Example 2.12. To determine whether $x_1 = x_2 = 0$ is a stable equilibrium of the system

$$\dot{x}_1 = (x_2 - 1)x_1^3$$

$$\dot{x}_2 = -\frac{x_1^4}{(1 + x_1^2)^2} - \frac{x_2}{1 + x_2^2} \tag{2.19}$$

using Lyapunov's direct method, we need to find a scalar function $V(x_1, x_2)$

satisfying some or all of the conditions of theorems 2.9 and 2.11. As a starting point, try the positive definite function

$$V(x_1, x_2) = x_1^2 + x_2^2.$$

Differentiating this function along system trajectories, we have

$$\dot{V}(x_1, x_2) = 2x_1 \dot{x}_1 + 2x_2 \dot{x}_2$$
$$= -2x_1^4 + 2x_2 x_1^4 - 2\frac{x_2 x_1^4}{(1 + x_1^2)^2} - 2\frac{x_2^2}{1 + x_2^2},$$

which is not of the required form since $\dot{V} \not\leq 0$ for some x_1, x_2. However it is not far off since the 2nd and 3rd terms in the above expression nearly cancel each other. After some experimentation with the first term in V we find that, with V redefined as

$$V(x_1, x_2) = \frac{x_1^2}{1 + x_1^2} + x_2^2, \tag{2.20}$$

(which is again positive definite) the derivative becomes

$$\dot{V}(x_1, x_2) = 2\left[\frac{x_1}{1 + x_1^2} - \frac{x_1^3}{(1 + x_1^2)^2}\right]\dot{x}_1 + 2x_2 \dot{x}_2$$
$$= -2\frac{x_1^4}{(1 + x_1^2)^2} - 2\frac{x_2^2}{1 + x_2^2},$$

so that $\dot{V}$ is now negative definite. Thus (2.20) satisfies all of the conditions of theorem 2.9, and the equilibrium is therefore asymptotically stable. However V is not radially unbounded since the contours of $V(x_1, x_2) = V_0$ are not closed curves for $V_0 \geq 1$ (figure 9), and it is therefore not possible to conclude that the system (2.19) is globally asymptotically stable. Neither is it possible to conclude that (2.19) is *not* globally asymptotically stable without further analysis, since some other function V might be found which both satisfies the conditions of theorem 2.9 and is radially unbounded. In fact (2.19) is not globally asymptotically stable since it is possible to find initial conditions for which $x_1(t), x_2(t)$ do not converge to zero (for example the trajectory starting from $(x_1, x_2) = (3, 1.5)$ shown in figure 9 escapes to infinity).   ◇

Exponential stability

Exponential stability is a special case of asymptotic stability in which the convergence of the system state to the equilibrium is bounded by an exponentially decaying function of time. To show that an equilibrium is exponentially stable using Lyapunov's direct method, it is therefore necessary to impose further conditions on the rate of change of a Lyapunov function which demonstrates asymptotic stability. The basic approach involves finding a scalar function $V(x,t)$ whose derivative along system trajectories satisfies

$$\dot{V}(x,t) \leq -aV(x,t) \tag{2.21}$$

for some constant $a > 0$, in a region of state space containing the equilibrium. Suppose therefore that (2.21) holds for all x such that $\|x\| \leq R_0$. If V is also positive definite and decrescent for $\|x\| \leq R_0$, then by theorem 2.10 there exists an $r > 0$ such that all trajectories $x(t)$ with initial conditions $x(t_0)$ satisfying $\|x(t_0)\| \leq r$ remain within the region on which $\|x\| \leq R_0$ for all $t \geq t_0$. On these trajectories the bound (2.21) holds at all times, which implies that

$$V\big(x(t),t\big) \leq V\big(x(t_0),t_0\big)e^{-a(t-t_0)}, \ \forall t \geq t_0. \tag{2.22}$$

whenever $\|x(t_0)\| \leq r$. In order to conclude that $\|x(t)\|$ satisfies a similar bound, we need to find a lower bound on $V(x,t)$ in terms of a positive power of $\|x\|$ of the form

$$V(x,t) \geq b\|x\|^p, \tag{2.23}$$

for some constants $b > 0$ and $p > 0$. Note that the positive definiteness of V ensures that it is always possible to construct such a bound for all x satisfying $\|x\| \leq R_0$. Combining inequalities (2.22) and (2.23) leads to the required result:

$$\|x(t_0)\| \leq r \implies \|x(t)\| \leq Re^{-\alpha t}, \ \forall t \geq t_0. \qquad \begin{cases} R = \left[\dfrac{V\big(x(t_0),t_0\big)}{b}\right]^{1/p} \\ \alpha = a/p \end{cases}$$

Clearly this argument can be used to show global exponential stability if R_0 can be taken to be infinite and V is radially unbounded. A variation on the above

approach assumes that $-\dot{V}$ is greater than some positive power of $\|x\|$, and then imposes both upper and lower bounds on V in terms of positive powers of $\|x\|$ in order to assert exponential stability.

Concluding remarks

- Lyapunov's direct method only gives sufficient conditions for stability. This means for example that the existence of a positive definite V for which $\dot{V} \not\preceq 0$ along system trajectories does not necessarily imply that $x = 0$ is unstable.

- Lyapunov analysis can be used to find conditions for instability of an equilibrium point, for example $x = 0$ must be unstable if V and $\dot{V}$ are both positive definite. These results are known as **instability theorems**.

- It is also possible to show that the premises and conclusions of each of theorems 2.9–2.11 are interchangeable, so for example there must be some positive definite function V with $\dot{V}$ negative definite if $x = 0$ is asymptotically stable. These results are known as **converse theorems**.

- There are no general guidelines for constructing Lyapunov functions. In situations where physical insights into the system dynamics do not suggest an obvious choice for V, it is often necessary to resort to a Lyapunov-like analysis based on a function satisfying only some of the conditions of theorems 2.9–2.11. This approach will be discussed in the next section.

3 Lyapunov-like convergence analysis [11]

A common problem with Lyapunov's direct method is that it allows asymptotic stability to be concluded only under extremely restrictive conditions. Thus it may be easy to find a positive definite function V which is non-increasing along system trajectories, but finding a V for which $\dot{V}$ is negative definite along system trajectories is usually a much harder task. However it is often possible to determine whether the system state, or at least some component of the state, converges asymptotically using a Lyapunov-like analysis based on a non-positive (rather than negative definite) derivative $\dot{V}$.

The techniques for Lyapunov-like convergence analyses are slightly different for autonomous and non-autonomous dynamics. This section first considers the case of non-autonomous systems of the form

$$\dot{x} = f(x, t), \qquad f(0, t) = 0, \ \forall t. \tag{3.1}$$

The method for autonomous systems, though more powerful and closer in principle to Lyapunov's direct method, will then be treated as a special case of that for non-autonomous systems.

3.1 Convergence of non-autonomous systems

Suppose that we have found a positive definite function $V(x, t)$ whose derivative $\dot{V}$ along trajectories of (3.1) satisfies

$$\dot{V}(x, t) \leq -W(x) \leq 0 \tag{3.2}$$

for some non-negative function W. The aim of the convergence analysis is to show that $W(x(t))$ tends to zero, and therefore that $x(t)$ converges to the set of states on which $W(x) = 0$. Specifically, (3.2) implies that V cannot increase along system trajectories, and since $V \geq 0$ due to the assumption that V is positive definite, it follows that $V(x(t), t)$ tends to a finite limit as $t \to \infty$:

$$0 \leq \lim_{t \to \infty} V(x(t), t) \leq V(x(t_0), t_0),$$

---

[11]See *Slotine and Li* §3.4.3 pp68–76 and §4.5 pp122–126 or *Vidyasagar* §5.3 pp176–186

where $x(t)$ is a trajectory of (3.1) with initial condition $x(t_0)$. Integration of (3.2) therefore yields a finite bound on the integral of $W(x(t))$ over the interval $t_0 \le t < \infty$,

$$\int_{t_0}^{\infty} W(x(t)) \, dt \le V(x(t_0), t_0) - \lim_{t \to \infty} V(x(t), t). \qquad (3.3)$$

Under certain conditions on W this bound leads to the conclusion that $W(x(t))$ converges to zero as t tends to ∞.

At first sight it may seem obvious that a non-negative function $\phi(t)$, which has finite integral over the infinite interval $0 \le t < \infty$, necessarily converges to zero. However this is not true in general. For example if ϕ were allowed to be discontinuous, then it would be possible to change the value of ϕ at any individual point t without affecting the integral of ϕ. In fact continuity is not enough to ensure convergence of ϕ, since it is possible to construct functions which are continuous at any finite time t, but which effectively become discontinuous as $t \to \infty$. Figure 10 gives an example of such a function; here $\phi(t)$ does not converge to a limit as $t \to \infty$ even though ϕ is continuous and the integral $\int_0^{\infty} \phi(t) \, dt$ is finite. A condition which does guarantee that $\phi(t) \to 0$ given that $\int_0^t \phi(s) \, ds$ converges to a finite limit as $t \to \infty$ is provided by a technical result known as Barbalat's lemma, which is stated here in a slightly simplified form.

Barbalat's lemma. For any function $\phi(t)$, if
- (a). $\dot{\phi}(t)$ exists and is finite for all t
- (b). $\lim_{t \to \infty} \int_0^t \phi(s) \, ds$ exists and is finite

then $\lim_{t \to \infty} \phi(t) = 0$.

From (3.3) it can therefore be concluded that $W(x(t))$ converges to zero asymptotically as $t \to \infty$ provided the derivative $\dot{W}$ of W along trajectories of (3.1) remains finite at all times t. Using the chain rule we have

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

and $\dot{W}(x(t))$ must therefore remain finite if W and f are continuous with respect to their arguments and $x(t)$ is bounded for all t.
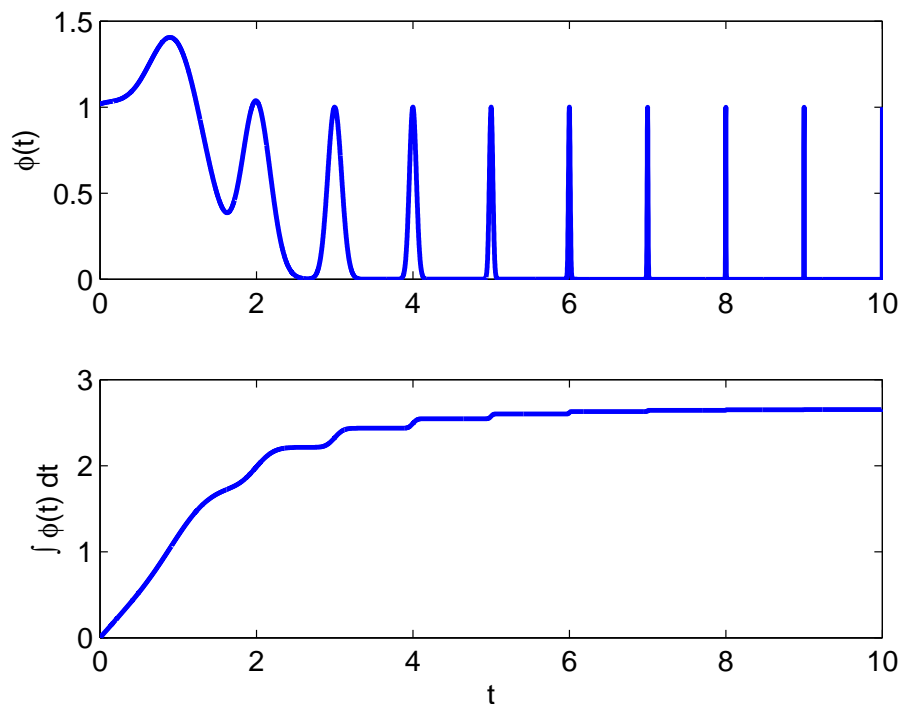
Figure 10: Example of a non-negative continuous function which does not converge to zero even though its integral tends to a finite limit as $t \to \infty$. Upper plot: $\phi(t) = \sum_{k=0}^{\infty} e^{-4^k(t-k)^2}$, lower plot: $\int_0^t \phi(s)\,ds$.

The following theorem summarizes this argument. For convenience the theorem assumes V to be decrescent and radially unbounded in order to ensure that $x(t)$ remains finite for all $t \geq t_0$ given arbitrary initial conditions $x(t_0)$. Clearly these additional assumptions are not needed if the boundedness of $x(t)$ is ascertained via a Lyapunov analysis based on an alternative Lyapunov function.

Theorem 3.1 (Convergence for non-autonomous systems)**.** *Let the function f in (3.1) be continuous with respect to x and t, and assume that there exists a continuously differentiable scalar function $V(x,t)$ such that:*

   (a). $V(x,t)$ is positive definite, radially unbounded and decrescent
   (b). $\dot{V}(x,t) \leq -W(x) \leq 0$

where W is a continuous function. Then all state trajectories $x(t)$ of (3.1) are globally bounded, and satisfy $W\big(x(t)\big) \to 0$ as $t \to \infty$.

3.2 Convergence of autonomous systems

The preceding analysis can be used to derive stronger convergence properties when the system dynamics are autonomous, ie. of the form

$$\dot{x} = f(x), \qquad f(0) = 0. \tag{3.4}$$

Essentially this is because it is much easier to determine whether the system state remains within a given region of state space if the system is autonomous rather than non-autonomous. Consequently an estimate of a set of points, say R, to which the state $x(t)$ of (3.4) converges can be refined using the criterion that, having entered R, $x(t)$ must remain within R at all future times.

To illustrate the approach, consider again the mass-spring-damper example of section 2.3. We have already shown that the equilibrium $y = \dot{y} = 0$ is stable using the argument of theorem 2.9. But it is not possible to show asymptotic stability by applying Lyapunov's direct method to the function V defined in (2.11), since the derivative $\dot{V} = -c(\dot{y})\dot{y}$ is not a negative definite function of the system state. However the stability of the equilibrium ensures that y and $\dot{y}$ remain finite on any system trajectory with initial conditions sufficiently close to the equilibrium. Application of Barbalat's lemma to $\dot{V}$ therefore shows that such trajectories converge to the set of states on which $\dot{y} = 0$. But if $\dot{y} = 0$, the acceleration $\ddot{y}$ is non-zero whenever the displacement y is non-zero, and consequently the system state cannot remain indefinitely at any point for which $y \neq 0$. This implies that the state must converge to the set on which $\dot{y} = 0$ *and* $y = 0$, ie. the origin. Thus the equilibrium $y = \dot{y} = 0$ is asymptotically stable.

This argument is based on the concept of an invariant set.

> **Definition 3.2** (Invariant set)**.** A set M is an **invariant set** for a dynamic system if every system trajectory starting in M remains in M at all future times.

Thus an invariant set has the same properties as an equilibrium point but, unlike an equilibrium, can consist of more than just a single point. Useful

examples of invariant sets are equilibrium points and limit cycles (ie. system trajectories which form closed curves in state space).

The method used above to show the asymptotic stability of the mass-spring-damper system can be generalized simply by noting that every bounded trajectory of an autonomous system converges to an invariant set. If it can be shown, using Barbalat's lemma for example, that the trajectories $x(t)$ of (3.4) converge to some set R, then it follows that $x(t)$ must converge to an invariant set M contained in R. The following theorem uses this observation to refine theorem 3.1, adapted for the case of autonomous dynamics.

Theorem 3.3 (Invariant set theorem). *Let f in (3.4) be continuous, and assume that there exists a continuously differentiable scalar function $V(x)$ such that:*

    (a). $V(x)$ is positive definite and radially unbounded

    (b). $\dot{V}(x) \leq 0$

then all solutions $x(t)$ of (3.4) are globally bounded. Furthermore, let R be the set of all x for which $\dot{V}(x) = 0$, and let M be the largest invariant set in R. Then every state trajectory $x(t)$ of (3.4) converges to M as $t \to \infty$.

Note that:

1. The set M is defined in theorem 3.3 as the largest invariant set in R in the sense that M is the union of all invariant sets (eg. equilibrium points or limit cycles) within R.

2. The Lyapunov-like convergence theorems of this section contain the global asymptotic convergence results of Lyapunov's direct method (theorem 2.11) as special cases; namely when $W(x)$ in the convergence theorem for non-autonomous systems is positive definite, and when M in the invariant set theorem consists only of the origin. In fact theorem 3.3 presents a significant generalization of Lyapunov's direct method by providing criteria for convergence to entire state trajectories such as limit cycles.

The region of attraction of the set M in theorem 3.3 is clearly the entire state space, since the conditions on V are required to hold for all x. However, by

relaxing slightly the conditions of theorem 3.3, a more general local invariant set theorem can be derived for autonomous systems. This is useful for determining regions of attraction for systems whose stability properties are local rather than global.

Theorem 3.4 (Local invariant set theorem). *Let f in (3.4) be continuous, and assume that there exists a continuously differentiable scalar function $V(x)$ such that:*

(a). for some constant $\overline{V} > 0$, the set Ω defined by $V(x) < \overline{V}$ is bounded

(b). $\dot{V}(x) \leq 0$ for all x in Ω

then Ω is an invariant set for (3.4). Furthermore, let R be the set of all points x in Ω for which $\dot{V}(x) = 0$, and let M be the largest invariant set in R. Then every solution $x(t)$ of (3.4) with initial conditions in Ω converges to M as $t \to \infty$.

To show that Ω is an invariant set, note that condition (b) of the theorem implies that $V(x(t))$ cannot exceed the value $\overline{V}$ along any trajectory starting in Ω. Convergence of $x(t)$ to M can be shown by applying Barbalat's lemma to $\dot{V}(x(t))$ along trajectories with initial conditions in Ω, and then invoking the condition that $x(t)$ must converge to an invariant set.[12]

Note that:

1. The local invariant set theorem does away with the requirement that V is positive definite by requiring instead that the level set Ω is bounded. This condition performs a similar function to the positive definite condition of Lyapunov's direct method since it ensures that the continuous function V is lower bounded on Ω (ie. $V(x) \geq \underline{V}$ for all x in Ω and some finite $\underline{V}$), and that any trajectory contained in Ω is bounded.

2. The set Ω is a region of attraction of the set M, though not necessarily the largest region of attraction.

---

[12]Barbalat's lemma is applicable here since V and $\dot{V}$ are necessarily finite on trajectories in Ω due to the assumptions that V, $\dot{V}$ and f are continuous and Ω is bounded.

Example 3.5. Consider the system

$$\dot{x}_1 = -x_2 + x_1(x_1^2 + x_2^2 - 1)$$
$$\dot{x}_2 = x_1 + x_2(x_1^2 + x_2^2 - 1)$$

(3.5)

which has an equilibrium at $x_1 = x_2 = 0$. To determine the stability of this equilibrium point, define V as the function

$$V(x_1, x_2) = x_1^2 + x_2^2.$$

Then the derivative of V along trajectories of (3.5) is given by

$$\dot{V}(x_1, x_2) = 2(x_1^2 + x_2^2)(x_1^2 + x_2^2 - 1).$$

Let Ω be the set of points (x_1, x_2) satisfying $V(x_1, x_2) < 1$, then Ω is bounded (since Ω is a unit disc centred on the origin) and $\dot{V}(x_1, x_2) \leq 0$ for all x in Ω. Thus all conditions of the local invariant set theorem are satisfied, and since the subset of Ω on which $\dot{V}(x_1, x_2) = 0$ is simply the point $x_1 = x_2 = 0$, it follows that any trajectory starting within Ω converges to the origin. Therefore the origin is an asymptotically stable equilibrium and Ω is contained within its region of attraction.

Even though Ω is the largest region of attraction that can be determined with this choice of V (since $\dot{V}(x_1, x_2) \not\leq 0$ for some (x_1, x_2) in the set $V(x_1, x_2) < \overline{V}$ whenever $\overline{V} > 1$), it cannot be concluded without further analysis that Ω contains every point in the region of attraction of the origin. In this example however, Ω is in fact the largest region of attraction of the origin since the set Ω' defined by $V(x_1, x_2) = 1$ is an unstable limit cycle, ie. all trajectories starting from points arbitrarily close to Ω' either converge asymptotically to the origin or tend to infinity.

To prove this, first note that the first and second derivatives of $x_1^2 + x_2^2 - 1$ along system trajectories are zero at any point satisfying $x_1^2 + x_2^2 = 1$. On every state trajectory with initial conditions $x_1(t_0), x_2(t_0)$ satisfying $x_1^2(t_0) + x_2^2(t_0) = 1$ we therefore have $x_1^2(t) + x_2^2(t) = 1$ for all $t \geq t_0$, which implies that Ω' is a limit cycle of (3.5). Next consider the function

$$V'(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2,$$

which has derivative

$$\dot{V}'(x_1, x_2) = 4(x_1^2 + x_2^2)(x_1^2 + x_2^2 - 1)^2$$

along trajectories of (3.5). Since V' is minimized on the set Ω' and $\dot{V}'$ is positive whenever $x_1^2 + x_2^2 \neq 0$, it can be concluded that Ω' is an unstable limit cycle. $\Diamond$
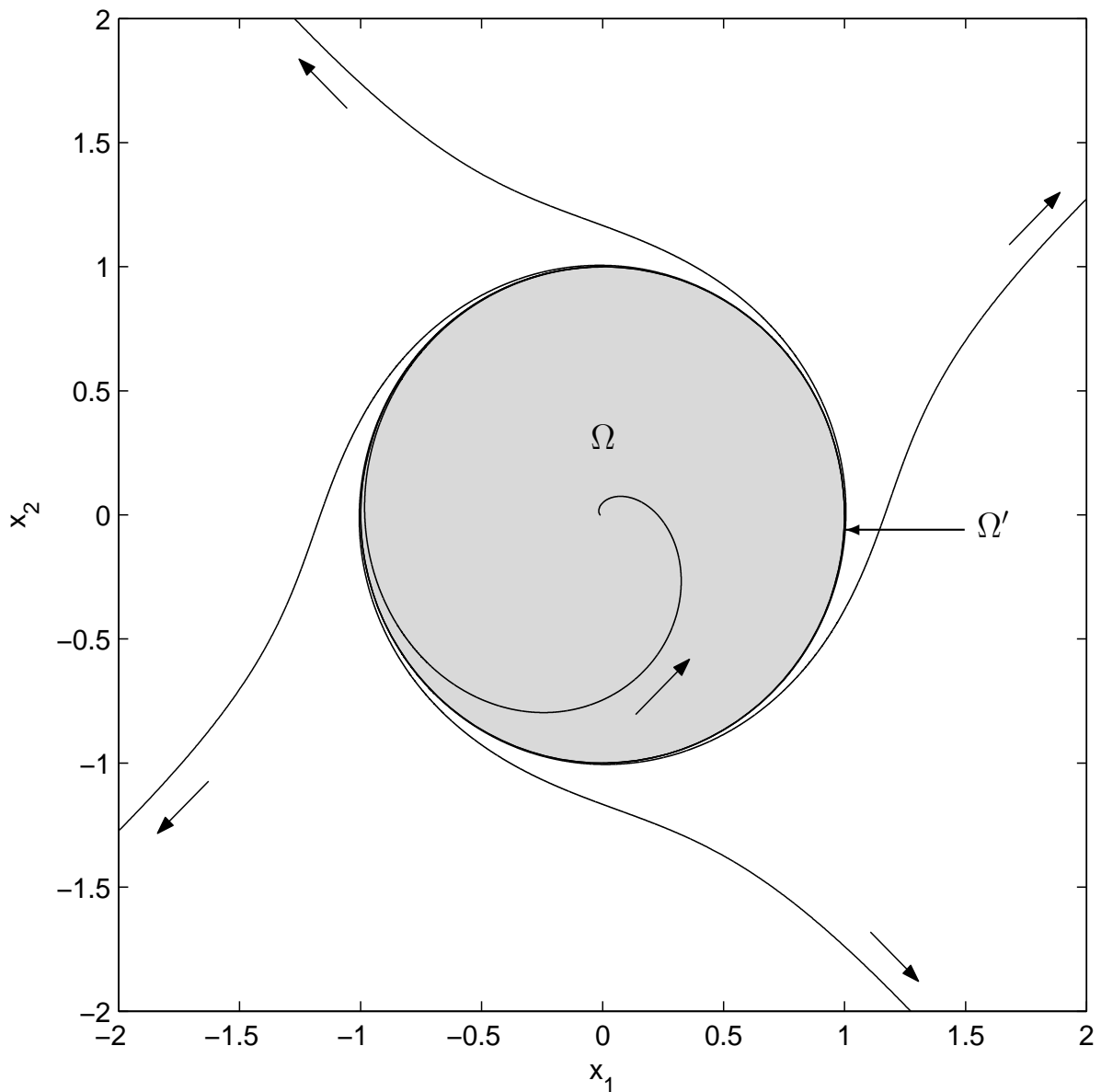

Figure 11: The region of attraction Ω and the unstable limit cycle Ω' in example 3.5.

4 Linear systems and passive systems

This section builds on the discussion of Lyapunov's direct method in the previous section by considering two classes of dynamic system, linear and passive systems. It is possible to find Lyapunov functions for these classes of system systematically, and the problem of determining Lyapunov functions for complex nonlinear systems is therefore simplified if linear or passive subsystems can be identified as components of the original dynamics. This section first considers the stability of linear systems within the Lyapunov stability framework, and then describes passivity and the properties of interconnected passive systems. It concludes with a discussion of the stability of feedback systems in which the forward path contains a linear subsystem and the feedback path contains a memoryless (though possibly time-varying) nonlinearity.

4.1 Linear systems [13]

The Lyapunov stability analysis of linear time-invariant (LTI) systems is based entirely on quadratic forms, i.e. functions of the form

$$V(x) = x^\top P x.$$

Before giving the details of the method, we first describe some properties of matrices and quadratic forms.

- Any square matrix P can be expressed as the sum of a symmetric matrix P_1 (satisfying $P_1 = P_1^\top$) and a skew-symmetric matrix P_2 (satisfying $P_2 = -P_2^\top$):

$$P = P_1 + P_2, \qquad \begin{cases} P_1 = \frac{1}{2}(P + P^\top), \\ P_2 = \frac{1}{2}(P - P^\top). \end{cases}$$

- If P_2 is skew-symmetric, then $x^\top P_2 x = x^\top P_2^\top x = -x^\top P_2 x$, and hence $x^\top P_2 x = 0$ for any vector x of conformal dimensions. Therefore a quadratic form $x^\top P x$ with non-symmetric P is equivalent to the quadratic form $\frac{1}{2} x^\top (P + P^\top) x$ involving the symmetric part of P alone.

---

[13]See *Slotine and Li* §3.5.1 pp77–83 or *Vidyasagar* §5.4.2 pp196–202

- A symmetric matrix P is **positive definite** (denoted $P \succ 0$) if the quadratic form involving P is a positive definite function, i.e. if

$$x^\top P x > 0 \text{ for all } x \neq 0.$$

  If $x^\top P x \geq 0$ for all $x \neq 0$, then P is **positive semidefinite** ($P \succeq 0$). Similarly, P is **negative definite** ($P \prec 0$) or **negative semidefinite** ($P \preceq 0$) respectively if $x^\top P x < 0$ or $x^\top P x \leq 0$ for all non-zero x.

- Any symmetric matrix P can be decomposed as

$$P = U \Lambda U^\top,$$

  where Λ is a diagonal matrix of eigenvalues of P, and U is an orthogonal matrix ($U^\top U = I$) containing the eigenvectors of P. Thus

$$x^\top P x = z^\top \Lambda z, \quad z = U^\top x,$$

  and it follows that P is positive definite if and only if all eigenvalues of P are strictly positive. By the same argument P is positive semidefinite, negative semidefinite, or negative definite if all eigenvalues of P are greater than or equal to zero, less than or equal to zero, or strictly negative, respectively.

To find a Lyapunov function for a given stable linear system

$$\dot{x} = Ax, \tag{4.1}$$

consider the positive definite function

$$V(x) = x^\top P x,$$

where P is some positive definite symmetric matrix (i.e. $P = P^\top \succ 0$). The derivative of V along trajectories of (4.1) is

$$\dot{V}(x) = x^\top (A^\top P + P A) x,$$

and the system is therefore globally asymptotically stable by Lyapunov's direct method if there exists a positive definite matrix Q satisfying the condition

$$A^\top P + P A = -Q. \tag{4.2}$$

This is known as a Lyapunov matrix equation.

For arbitrary positive definite P, the matrix Q in (4.2) will not necessarily be positive definite. However it is always possible to choose a matrix $Q = Q^\top \succ 0$ and solve (4.2) for $P = P^\top \succ 0$ if (4.1) all eigenvalues of A have strictly negative real parts. To see this, suppose that a particular Q has been chosen. Then along a trajectory of the system (4.1) with initial condition $x(t_0)$, the derivative of the quadratic form V defined in terms of a P satisfying the Lyapunov equation (4.2) is given by

$$\dot{V}\big(x(t)\big) = -x^\top(t_0)e^{A^\top(t-t_0)}Qe^{A(t-t_0)}x(t_0).$$

Since (4.1) is stable, $V\big(x(t)\big)$ must converge to zero as $t \to \infty$. The integral of $\dot{V}$ with respect to t therefore gives

$$x(t_0)^\top P x(t_0) = x^\top(t_0)\left[\int_0^\infty e^{A^\top t}Qe^{At}dt\right]x(t_0).$$

This must be true for any initial condition $x(t_0)$, so that

$$P = \int_0^\infty e^{A^\top t}Qe^{At}dt. \tag{4.3}$$

The integral on the RHS is well-defined due to the assumption that (4.1) is stable, and thus defines a positive definite matrix P satisfying the Lyapunov equation (4.2) for given positive definite Q.[14]

This discussion can be summarized as follows.

> **Theorem 4.1** (Lyapunov stability of LTI systems)**.** *A necessary and sufficient condition for a LTI system $\dot{x} = Ax$ to be asymptotically stable is that, for any positive definite symmetric matrix Q, the Lyapunov matrix equation (4.2) has a unique solution P which is a positive definite symmetric matrix.*

A Lyapunov function for the linear system (4.1) can therefore computed by choosing $Q \succ 0$ and solving (4.2) for P. The system (4.1) is stable if and only if $P \succ 0$. For given A and Q, the Lyapunov equation (4.2) is a set of linear

---

[14]The matrix P in (4.3) must be positive definite since $x^\top P x = 0$ implies that $x^\top e^{A^\top t}Qe^{At}x = 0$ for all $t \geq 0$, and hence $x^\top Q x = 0$, which implies that $x = 0$ since $Q \succ 0$.

equations in the elements of P of the form $(A^\top \otimes I + I \otimes A^\top)\text{vec}(P) = -\text{vec}(Q)$ (where $\otimes$ is the Kronecker product and $\text{vec}(\cdot)$ is the vector of stacked columns of a matrix), which can be solved for $\text{vec}(P)$ using a linear equation solver.

4.2 Passive systems [15]

Just as a Lyapunov function can be thought of as a generalization of the energy stored within a system, passivity generalizes the property of energy dissipatioxn. Passive systems have no internal sources of power, or, more generally, the rate of dissipation of energy in a passive system exceeds the rate of internal energy generation. Motivated by the principle of energy conservation in physical systems, this leads to the following definition of passivity for nonlinear systems of the form

$$\begin{aligned} \dot{x} &= f(x, u, t) \\ y &= h(x, t) \end{aligned} \qquad \begin{cases} u: & \text{system input,} \\ y: & \text{system output.} \end{cases} \qquad (4.4)$$

> **Definition 4.2** (Passivity). The system (4.4) is **passive** if there exists a continuous function $V(x, t)$ such that $V \geq 0$, $V(0, t) = 0$, and along system trajectories
> $$\dot{V}(x, t) = y(t)^\top u(t) - g(t) \qquad (4.5)$$
> for all t, for some function $g(t) \geq 0$.
> If $\int_0^t g \, dt \geq \epsilon \int_0^t u^\top u \, dt$ for some $\epsilon > 0$ and all t, then (4.4) is **strictly passive**.

The term $y^\top u$ in (4.5) corresponds to the net power input to the system, while g represents the rate of energy dissipation within the system[16].

Example 4.3. The system:

$$m\ddot{x} + x^2 \dot{x}^3 + x^7 = F$$

is a passive mapping (from F to $\dot{x}$) because

$$\frac{d}{dt}\left(\tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{8}x^8\right) = \dot{x}F - x^2\dot{x}^4. \qquad \diamond$$

---

[15] See *Slotine and Li* §4.7 pp132–142

[16] Strictly passive systems defined in Def. 4.2 are also called "input strictly passive" or "dissipative"

There are usually many different possible ways to define the output of a system. However it makes sense to choose as the output a signal which forms the input to another connected subsystem. This highlights the usefulness of the passivity property: it allows a Lyapunov function for a system of interconnected passive systems to be constructed from the sum of the functions V for each passive subsystem.

Consider for example the feedback connection of a pair of passive systems S_1 and S_2 shown in Figure 12. From the definition of passivity, there exist non-negative functions V_1 and V_2 satisfying

$$\dot{V}_1 = y_1^\top u_1 - g_1, \qquad g_1 \geq 0,$$
$$\dot{V}_2 = y_2^\top u_2 - g_2, \qquad g_2 \geq 0.$$

Therefore the function $V_1 + V_2$ has derivative

$$\dot{V}_1 + \dot{V}_2 \leq y_1^\top u_1 + y_2^\top u_2 - g_1 - g_2$$
$$= y_1^\top (u_1 + y_2) - g_1 - g_2 = -g_1 - g_2,$$

and applying the convergence analysis of section 3, it follows that $g_1(t)$ and $g_2(t)$ converge to zero as $t \to \infty$. Furthermore, if V_1 and V_2 are positive definite decrescent functions of the state of the systems S_1 and S_2 respectively, then $V_1 + V_2$ is a Lyapunov function which shows uniform stability of the equilibrium at $x = 0$ of the overall system.
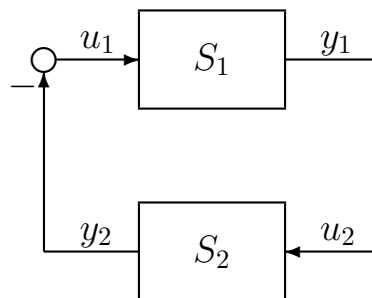

Figure 12: Feedback connection of passive systems.

Passive systems can form the building blocks of larger passive systems. In particular, both the feedback connection of Figure 13(a) and the parallel connection of Figure 13(b) result in passive dynamics mapping the overall input u to the overall output y whenever the subsystems S_1 and S_2 are both passive.

Again this can be shown by considering the sum of the storage functions V_1 and V_2 associated with the individual subsystems S_1 and S_2.
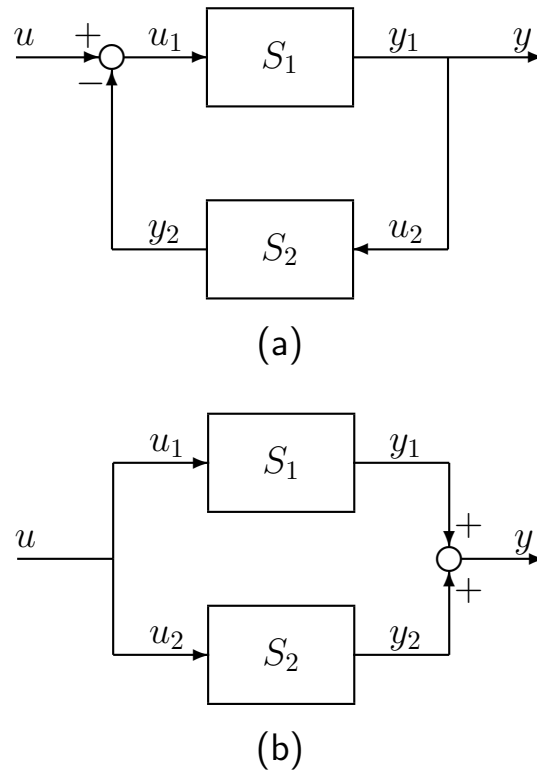

(a)

(b)

Figure 13: Interconnected passive subsystems. (a) Feedback connection. (b) Parallel connection.

Example 4.4 (Nonlinear adaptive control)**.** Suppose that we want to control the first order plant

$$\dot{x} = u + \theta x^2 \tag{4.6}$$

so that the state x converges to zero, where θ is an unknown but constant parameter. Let $\hat{\theta}$ be an estimate of θ which is allowed to vary over time according to the information on θ contained in measurements of x. The construction of the estimate $\hat{\theta}$ and the design of a control law based on this estimate is an adaptive control problem.

To solve this problem using a Lyapunov-based design method, consider the derivative of the function $V_1 = \frac{1}{2}x^2$ along system trajectories:

$$\dot{V}_1 = x(u + \theta x^2).$$

If the estimate $\hat{\theta}$ were exact (i.e. if $\hat{\theta} = \theta$), then the control law

$$u = -kx - \hat{\theta}x^2, \qquad k > 0, \tag{4.7}$$

would give $\dot{V}_1 = -kx^2$, and would therefore make $x = 0$ an asymptotically stable equilibrium point of the resulting closed-loop system.[17] However, for non-zero parameter estimate errors $\theta - \hat{\theta}$ the control law (4.7) may or may not be stabilizing, depending on the variation of $\hat{\theta}$ over time. With u defined as in (4.7), the derivative of V_1 is given by

$$\dot{V}_1 = (\theta - \hat{\theta})x^3 - kx^2.$$

Thus V_1 satisfies condition (4.5) of the passivity definition (with input $\theta - \hat{\theta}$, output x^3, and $g = kx^2$), which shows that this control law results in passive dynamics from the error $\theta - \hat{\theta}$ to x^3. It follows that both the estimation error and the equilibrium $x = 0$ of the closed-loop system can be stabilized by updating $\hat{\theta}$ so that the dynamics between x^3 and $-(\theta - \hat{\theta})$ are passive (Fig. 14). One such update law is

$$\dot{\hat{\theta}} = x^3 \tag{4.8}$$

since this ensures that the function $V_2 = \frac{1}{2}(\theta - \hat{\theta})^2$ has derivative (see Fig. 14).

$$\dot{V}_2 = -(\theta - \hat{\theta})\dot{\hat{\theta}} = -(\theta - \hat{\theta})x^3$$
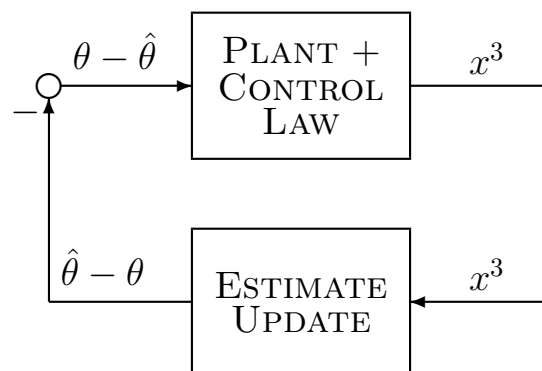

Figure 14: Feedback connection of the subsystems formed by the plant + control law and the estimate update law. Both are passive when the signals $\pm(\theta - \hat{\theta})$ and x^3 are considered as inputs and outputs.

To check that the combination of the control law (4.7) and parameter update law (4.8) meet the control objective, consider the function

$$V = V_1 + V_2 = \frac{1}{2}x^2 + \frac{1}{2}(\theta - \hat{\theta})^2.$$

[17]For this reason (4.7) is known as a *certainty equivalent* control law.

Along trajectories of the closed-loop system (4.6), (4.7), (4.8), we have

$$\dot{V} = \dot{V}_1 + \dot{V}_2 = -kx^2.$$

Therefore the estimate $\hat{\theta}(t)$ remains finite for all t and $x = 0$ is a globally asymptotically stable equilibrium of the closed-loop system.   ◊

Linear passive systems

For linear systems, passivity has a convenient interpretation in terms of the system frequency response. Consider the transfer function model

$$\frac{Y(s)}{U(s)} = H(s). \tag{4.9}$$

Theorem 4.5. *The system (4.9) is passive if and only if the poles p_i of H satisfy $\mathrm{Re}(p_i) \leq 0$ and the real part of its frequency response function is non-negative:*

$$\mathrm{Re}\big[H(j\omega)\big] \geq 0, \ \text{for all } \omega \geq 0. \tag{4.10}$$

This is easy to show using Parseval's theorem, which implies that the integral of the product $y(t)u(t)$ is non-negative if and only if (4.10) is satisfied. Similarly (4.9) is strictly passive[18] if its poles lie strictly in the left half plane and

$$\mathrm{Re}\big[H(j\omega)\big] > 0, \ \text{for all } \omega \geq 0. \tag{4.11}$$

The **Kalman-Yakubovich-Popov lemma** relates the frequency response condition (4.11) to the state space of (4.9). A simplified version is given next.

Lemma 4.6 (Kalman-Yakubovich-Popov). *If (4.9) is strictly passive, then there exist positive definite matrices P and Q such that* [19]

$$V(x) = x^\top P x, \qquad \dot{V}(x) = yu - x^\top Q x$$

where x is the state (of any controllable and observable realization) of (4.9).

---

[18]Passive and strictly passive linear systems are sometimes referred to as positive real systems and strictly positive real systems respectively.

Despite the useful properties of passive systems, it must be recognized that passivity is quite a restrictive conditions for the following reasons.

- A passive system must be open-loop stable (i.e. stable for $u(t) = 0$).

- The relative degree (number of poles minus zeros) of a passive system must be 0 or 1.

The condition on the relative degree is a direct consequence of the passivity condition (4.10). However, it also applies more generally to nonlinear passive systems, for which the relative degree is defined as the number of times the output must be differentiated before an expression containing the input is obtained.

4.3 Linear systems and nonlinear feedback[20]

This section considers the stability of the class of feedback systems with the structure shown in Figure 15. Here the forward path contains a linear time-invariant system H, and the feedback is via a nonlinear function ϕ. We assume that the input and output of H are scalars (multivariable generalisations exist). The dynamics of the linear system H are given by

$$\dot{x} = Ax + bu \qquad y = c^\top x \tag{4.12}$$

where (A, b, c) is a state-space realization of the transfer function $H(s) = c^\top(sI - A)^{-1}b$, and the feedback path is specified by

$$u = -z \qquad z = \phi(y). \tag{4.13}$$

The nonlinearity ϕ is memoryless (i.e. the mapping $z = \phi(y)$ contains no dynamics), but is allowed to be time-varying. It is assumed that ϕ satisfies a sector condition, defined as follows.

Definition 4.7. A continuous function ϕ belongs to the sector $[a, b]$ if there exist two numbers a and b such that

$$a \leq \frac{\phi(y)}{y} \leq b \tag{4.14}$$

whenever $y \neq 0$, and $\phi(0) = 0$.

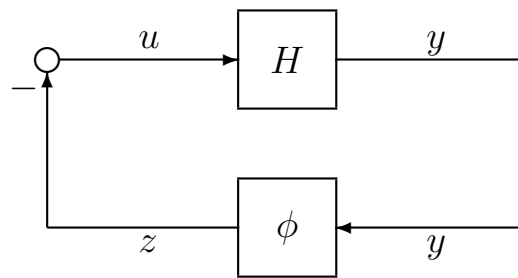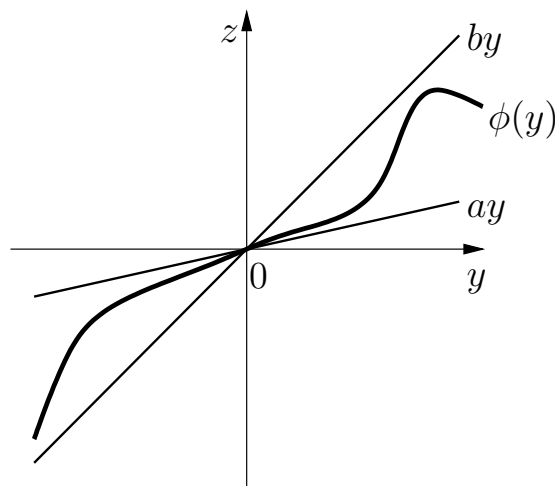[20]See *Slotine and Li* §4.8 pp142–147, or *Khalil* §10.1 pp400–418.

Figure 15: System structure.

The graphical interpretation of the sector condition (4.14) is simply that $\phi(y)$ lies between the two lines ay and by, as shown in Figure 16.

Figure 16: A function ϕ belonging to the sector $[a, b]$.

Systems of this form are of considerable practical interest and a number of different criteria for their stability have been determined. We will derive one of these, known as the circle criterion. This criterion has a graphical interpretation in terms of the frequency response $H(j\omega)$ which generalizes the Nyquist criterion for stability of linear feedback systems.

The circle criterion is based on the concept of passivity introduced in section 4.2. If the linear system H is strictly passive, i.e. if all poles of H have strictly negative real part and the frequency response $H(j\omega)$ satisfies condition (4.11), then the closed-loop system is guaranteed to be globally asymptotically stable whenever ϕ belongs to the sector $[0, \infty)$. This follows from the Kalman-Yakubovich-Popov lemma and the fact that $\phi(y)$ then has the same

sign as y. Specifically, the Kalman-Yakubovich-Popov lemma ensures that there exists a positive definite symmetric matrix P such that the function

$$V(x) = x^\top P x$$

has derivative

$$\dot{V}(x) = yu - x^\top Q x,$$

for some positive definite Q, along trajectories $x(t)$ of (4.12). The feedback control law (4.13) therefore gives

$$\dot{V}(x) = -y\phi(y) - x^\top Q x$$

where $y\phi(y) \geq 0$ for all y due to the assumption that ϕ belongs to the sector $[0, \infty)$. Therefore

$$\dot{V}(x) \leq -x^\top Q x$$

which implies that the equilibrium $x = 0$ is globally asymptotically stable by Lyapunov's direct method.

This argument can be extended to cover cases in which H is not passive and ϕ belongs to a general sector $[a, b]$ using loop transformations. The idea is to construct an equivalent closed-loop system in which the feedback path contains a nonlinearity belonging to the sector $[0, \infty)$ by adding feedforward and feedback loops to the subsystems of Figure 15. This approach allows the preceding argument to be used when ϕ lies outside the sector $[0, \infty)$ by placing more severe restrictions than (4.10) on the frequency response $H(j\omega)$. Alternatively it enables the argument to be applied when H is open-loop unstable by exploiting more precise information on the sector to which ϕ belongs.

Figures 17 and 18 show the two basic types of loop transformation. The signals u, y and z in each of these closed-loop systems are identical to those in the original system in figure 15, given the same initial conditions for the system H. For example the input to H in Figure 17 is given by

$$u = u' - ky = -z' - ky = -(z - ky) - ky = -\phi(y),$$

Similarly, the input to H in Figure 18 is

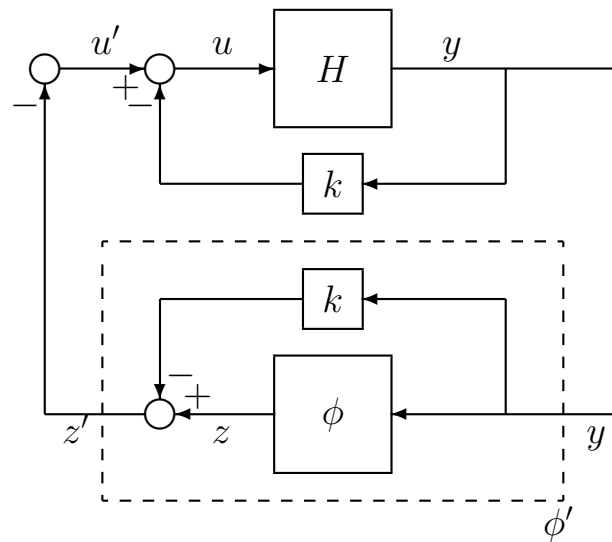$$u = -z = -\phi(y' + kz) = -\phi\big((y + ku) + kz\big) = -\phi(y).$$

Figure 17: Loop transformation of the system in Figure 15. If ϕ belongs to the sector $[a, b]$ then ϕ' belongs to the sector $[a - k, b - k]$.
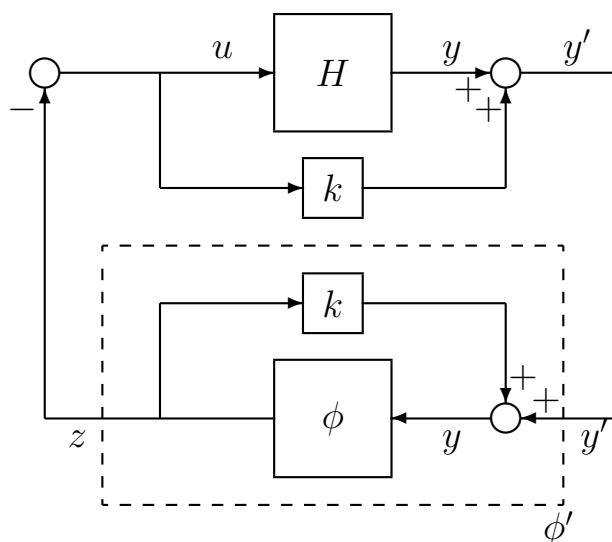

Figure 18: Loop transformation of the system in Figure 15. If ϕ belongs to the sector $[a, b]$ then ϕ' belongs to the sector $[a/(1 - ak), b/(1 - bk)]$.

However the mapping ϕ' between y and z' in Figure 17 is given by

$$z' = \phi'(y)$$
$$= \phi(y) - ky$$

so ϕ' belongs to the sector $[a - k, b - k]$ whenever ϕ belongs to $[a, b]$. On the other hand the mapping ϕ' between y' and z in Figure 18 satisfies

$$z = \phi'(y') \iff y' = \phi'^{-1}(z)$$
$$= \phi^{-1}(z) - kz.$$

If ϕ belongs to the sector $[a, b]$ (so that the inverse mapping ϕ^{-1} belongs to $[1/b, 1/a]$), then the inverse of ϕ' therefore lies in the sector $[1/b - k, 1/a - k]$, which implies that ϕ' belongs to $[a/(1 - ak), b/(1 - bk)]$.

Both types of loop transformation are employed in Figure 19 in order to construct a nonlinearity ϕ' which belongs to the sector $[0, \infty)$ whenever the original nonlinear function ϕ belongs to $[a, b]$. The Kalman-Yakubovich-Popov lemma can therefore be used to show that the closed-loop system in Figure 15 is asymptotically stable provided the transformed linear subsystem H' in Figure 19 is (i) strictly stable and (ii) satisfies $\mathrm{Re}[H'(j\omega)] > 0$ for all $\omega \geq 0$.

The subsystem H' contains a negative feedback loop around H with gain a. Consequently the stability of H' can be determined via the familiar Nyquist criterion for closed-loop stability. Thus if $H(s)$ has ν poles with positive real part, then all poles of $H'(s)$ have strictly negative real part if and only if the plot of $H(j\omega)$ makes ν anti-clockwise encirclements of the point $-1/a$ as ω goes from $-\infty$ to $+\infty$.

The condition on the real part of $H'(j\omega)$ has the following graphical interpretation. The frequency response of H' is given by

$$H'(j\omega) = \frac{H(j\omega)}{1 + aH(j\omega)} + \frac{1}{b - a}$$

and routine algebra shows that $\mathrm{Re}[H'(j\omega)] > 0$ for all $\omega \geq 0$ if and only if

$$\left| H(j\omega) + \tfrac{1}{2}(1/a + 1/b) \right|^2 > \tfrac{1}{4}(1/a - 1/b)^2, \quad \text{if } ab > 0 \qquad (4.15a)$$

$$\left| H(j\omega) + \tfrac{1}{2}(1/a + 1/b) \right|^2 < \tfrac{1}{4}(1/a - 1/b)^2, \quad \text{if } ab < 0 \qquad (4.15b)$$
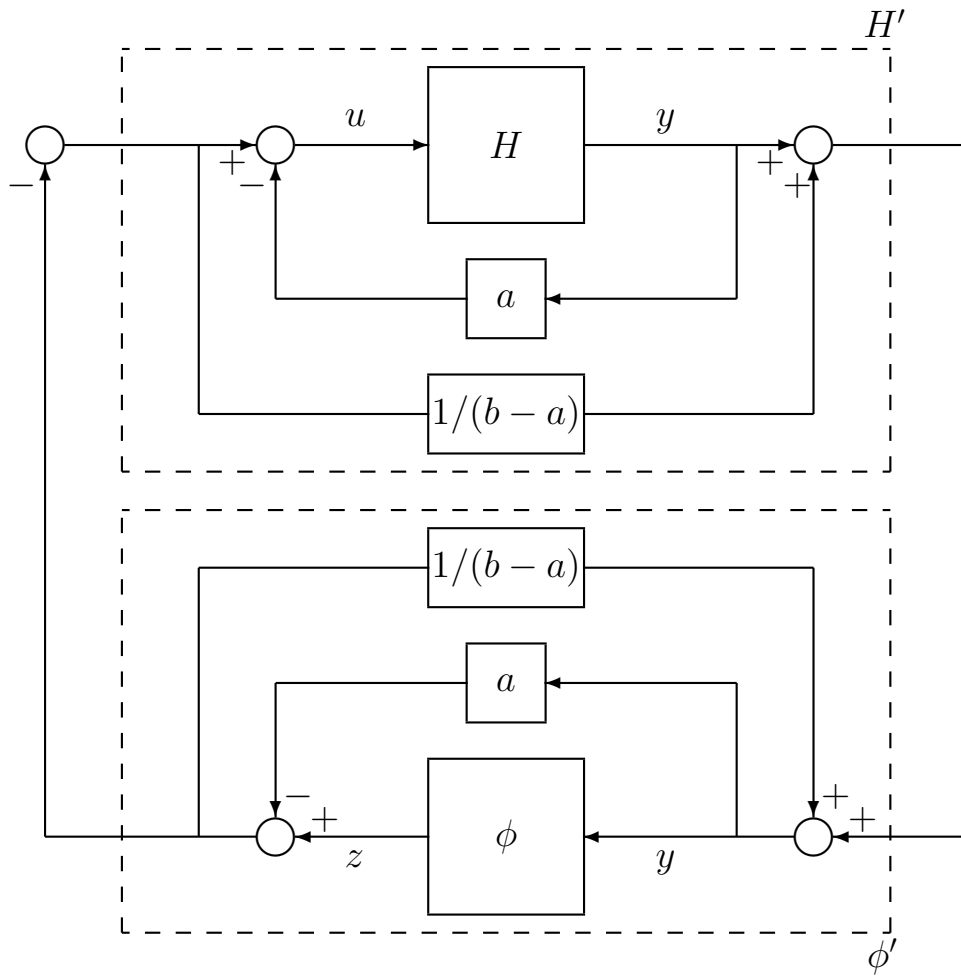
Figure 19: Loop transformation of the system in Figure 15. If ϕ belongs to the sector $[a, b]$ then ϕ' belongs to the sector $[0, \infty)$.

for all $\omega \geq 0$. Let $D(a, b)$ denote the disc in the complex plane with centre $-\frac{1}{2}(1/a+1/b)$ and radius $\frac{1}{2}(1/a-1/b)$ (see Figure 20). Then condition (4.15a) is satisfied for all $\omega \geq 0$ if the plot of $H(j\omega)$ does not enter the disc $D(a, b)$, while condition (4.15b) is satisfied if the plot of $H(j\omega)$ remains inside the disc $D(a, b)$ for all $\omega \geq 0$.
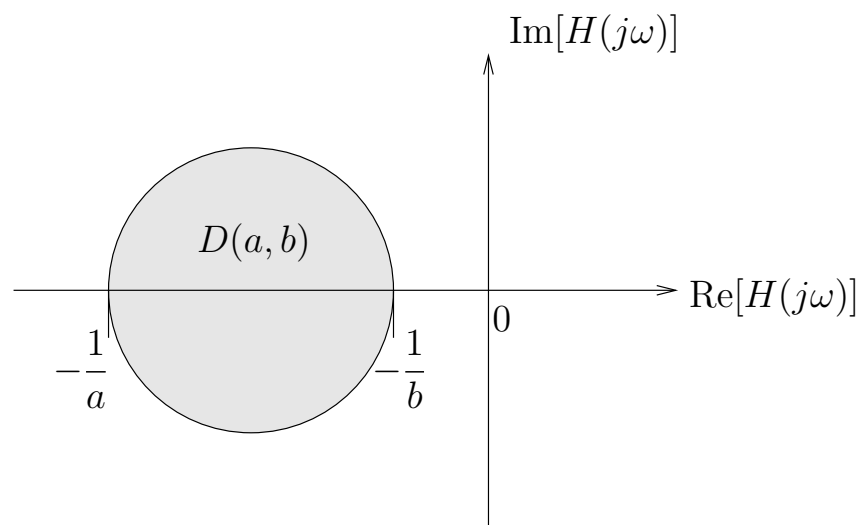
Figure 20: The disc $D(a, b)$ in the circle criterion.

These conditions are summarized in the following statement of the circle criterion.

Theorem 4.8 (Circle Criterion). *If the system (4.12)-(4.13) satisfies the conditions:*

    (a). $H(s)$ has ν poles with positive real part

    (b). the nonlinearity ϕ belongs to the sector $[a, b]$

    (c). one of the following is true

- *$0 < a < b$, and the Nyquist plot of $H(j\omega)$ does not enter the disc $D(a, b)$ and encircles it ν times anti-clockwise*
- *$0 = a < b$, $\nu = 0$, and the Nyquist plot of $H(j\omega)$ stays in the half-plane $\mathrm{Re}(s) > -1/b$*
- *$a < 0 < b$, $\nu = 0$, and the Nyquist plot of $H(j\omega)$ stays in the interior of the disc $D(a, b)$*
- *$a < b < 0$, and the Nyquist plot of $-H(j\omega)$ does not enter the disc $D(-a, -b)$ and encircles it ν times anti-clockwise*

then the equilibrium point $x = 0$ of the closed-loop system is globally asymptotically stable.

Note that:

1. The circle criterion allows the stability of the nonlinear system to be determined from the frequency response of the linear subsystem, which is relatively easy to compute experimentally.

2. The disc $D(a, b)$ performs roughly the same function in the circle criterion as the critical point $-1/a$ in the Nyquist criterion. As $b \to a$, the sector bound on ϕ shrinks, and the disc $D(a, b)$ tends towards the critical point $-1/a$.

3. Unlike the Nyquist criterion, the circle criterion gives sufficient (and not necessary) conditions for stability.

Example 4.9. A nonlinear system has the structure shown in Figure 15, where ϕ is a nonlinear function belonging to some sector $[a, b]$, and H is a linear time-invariant system with transfer function

$$H(s) = \frac{10}{(s-1)(s+3)^2}.$$

Since $H(s)$ has one pole in the right half complex plane (at $s = 1$), the Nyquist plot of $H(j\omega)$ must encircle the disc $D(a, b)$ anti-clockwise once in order to ensure stability by the circle criterion. From the plot of $H(j\omega)$ in Figure 21, this condition requires that a and b are both positive and is satisfied for example with $a = 1$, $b = 1.3$. $\diamond$
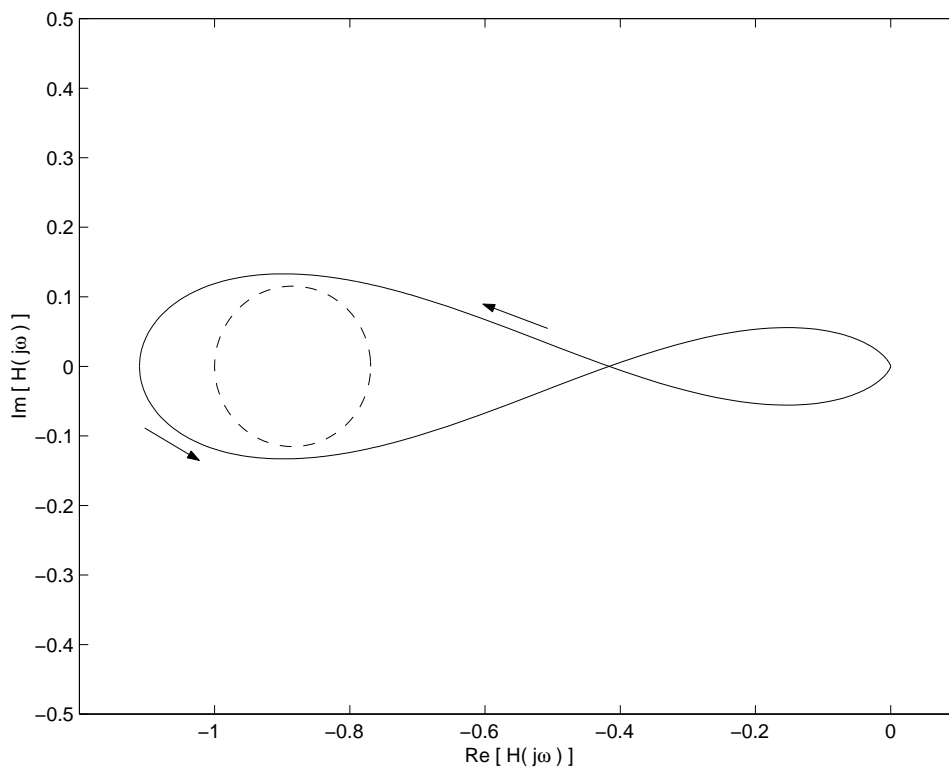

Figure 21: Nyquist plot of $H(j\omega)$ in example 4.9 (solid line) and the disc $D(1, 1.3)$ (dashed line).

4.4   Design example: Hydraulic active suspension system

An active suspension system for a train carriage uses a hydraulic actuator in series with mechanical springs and oil-filled dampers, arranged as shown in Figure 22 for a single carriage wheel. The purpose of the hydraulic actuator

is to reduce tilting caused by slowly varying loads acting on the carriage body (e.g. while cornering). The inclusion of the hydraulic actuator therefore allows for the use of a softer spring-damper assembly, resulting in better vibration isolation and a smoother ride.
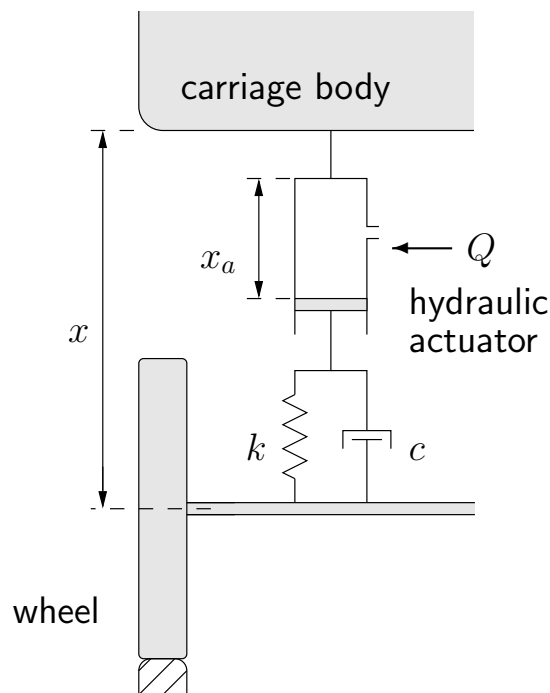

Figure 22: Active suspension for a train carriage.

The flow Q of fluid into the hydraulic actuator and its extension x_a (assuming incompressible fluid) are determined from

$$Q = \phi(u)$$
$$\dot{x}_a = Q/A$$

where u is the valve control signal, ϕ is a nonlinear function giving the valve characteristics, and A is the working area of the actuator. The function ϕ is not known exactly because of fluctuations in the pressure of the hydraulic fluid, but sector bounds on ϕ are available:

$$\phi \in [0.005, 0.1].$$

A controller for u is to be designed in order to compensate for the effects of unknown constant (or slowly varying) disturbance loads on the displacement x of the carriage body despite the uncertain valve characteristics $\phi(u)$.

The force exerted on the carriage by the suspension unit is F_susp:

$$F_\text{susp} = k(x_a - x) + c(\dot{x}_a - \dot{x})$$
$$= \left(k \int^t Q\,dt + cQ\right)/A - kx - c\dot{x}, \qquad Q = \phi(u)$$

If F and m are respectively the disturbance load that is to be rejected and the effective carriage mass acting on the suspension, then the dynamic model of carriage plus suspension is given by

$$F_\text{susp} - F = m\ddot{x}$$
$$\implies m\ddot{x} + c\dot{x} + kx = \left(k \int^t Q\,dt + cQ\right)/A - F, \qquad Q = \phi(u),$$

or, in terms of transfer functions

$$X(s) = \frac{cs + k}{ms^2 + cs + k} \cdot \frac{Q(s)}{As} - \frac{F}{ms^2 + cs + k}, \qquad Q = \phi(u).$$

Therefore, defining the control signal u feedback via a compensator with transfer fn. $C(s)$:

$$U(s) = C(s)E(s), \quad e = -x, \quad \text{setpoint: } x = 0$$

leads to a closed-loop system with the block diagram of Figure 23.
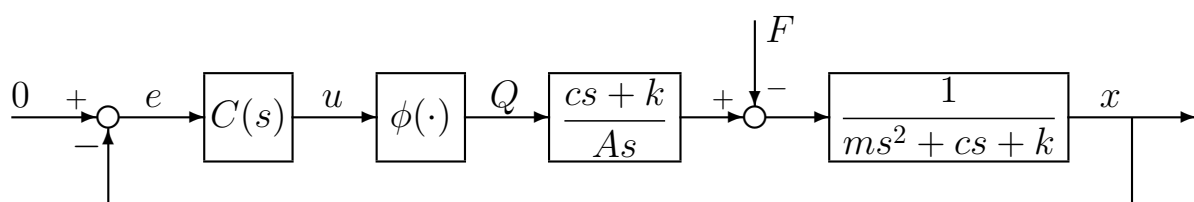

Figure 23: Closed-loop system with compensator $C(s)$.

Using $Q = \phi(u)$, $u = -C(s)X(s)$, this can be rearranged so that the nonlinearity is in the feedback path as shown in Fig. 24, which (for $F = 0$) now has the form of Fig. 15, with

$$H(s) = \frac{cs + k}{As(ms^2 + cs + k)}.C(s)$$
$$\phi \in [0.005, 0.1].$$

$$\frac{As}{cs+k} \cdot F \xrightarrow{+} \bigcirc \xrightarrow{} \boxed{\frac{cs+k}{As}} \xrightarrow{} \boxed{\frac{1}{ms^2+cs+k}} \xrightarrow{-x} \boxed{C(s)} \xrightarrow{u}$$
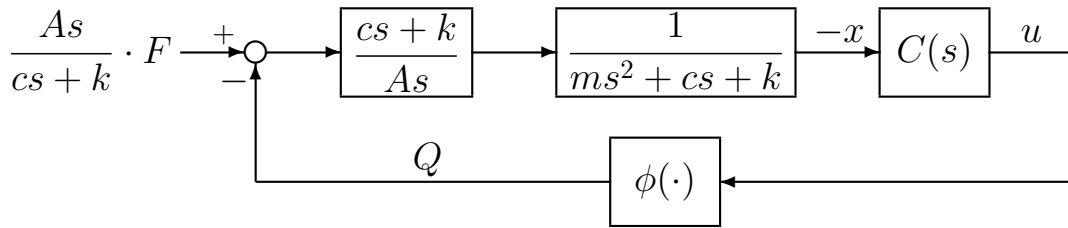

Figure 24: Closed-loop system.

To decide on the structure of $C(s)$, note that the hydraulics introduce an integrator into the forward path transfer function, so no integral term is needed in the controller in order to reject the constant disturbance F. Therefore consider a proportional + derivative compensator:

$$C(s) = K(1+\alpha s) \quad \Longrightarrow \quad H(s) = \frac{K(1+\alpha s)(cs+k)}{As(ms^2+cs+k)}$$

From the circle criterion, closed-loop (global asymptotic) stability is ensured if $H(j\omega)$ lies outside the disc $D(0.005, 0.1)$, for which a sufficient condition is that

$$\mathrm{Re}\big[H(j\omega)\big] > -10.$$

From the Nyquist plots of Figure 25, the value of α that maximizes the critical gain K and the corresponding value of K are given by

$$\alpha = 0.4$$
$$K \leq 10/3.4 = 2.94$$

Note that the maximum value for K is of interest here since the response of the closed-loop system is fastest when K is maximized.
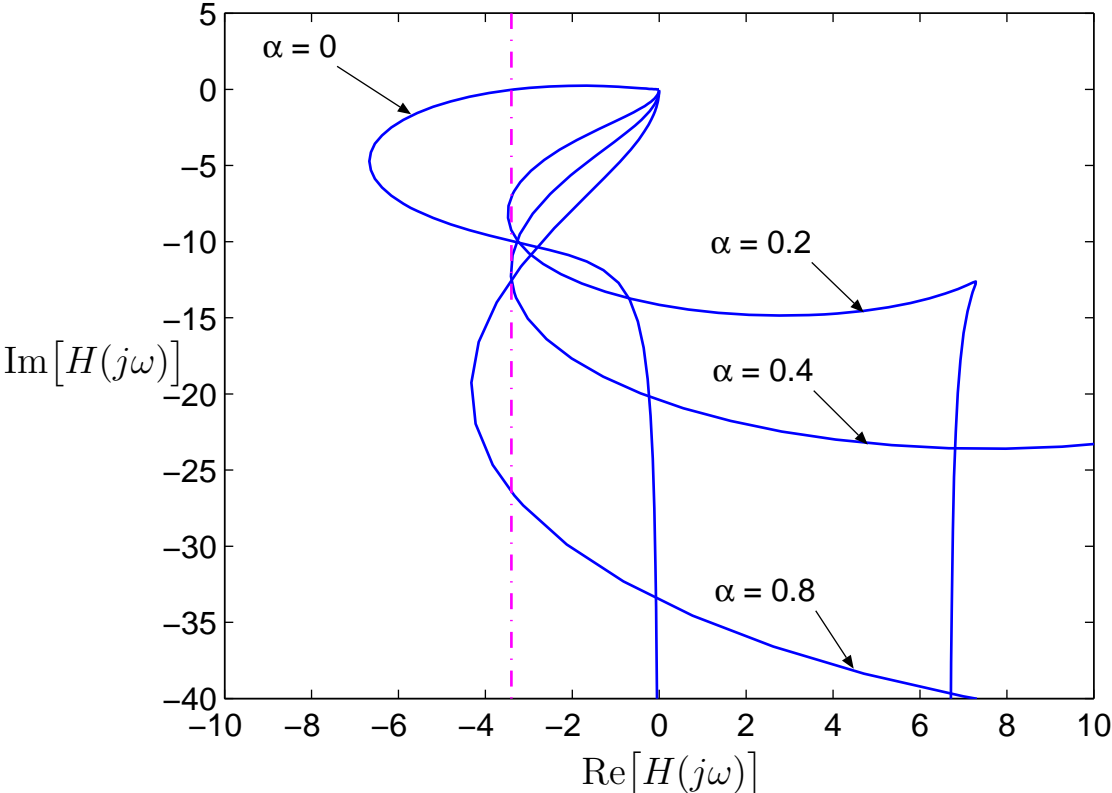
Figure 25: Nyquist plot of $H(j\omega)$ for $K = 1$ and $\alpha = 0, 0.2, 0.4, 0.8$