Chapter 2
Mathematical Preliminaries

2.1 Introduction

In this chapter, the fundamental mathematical concepts and analysis tools in systems theory are summarized, which will be used in control design and stability analysis in the subsequent chapters. Much of the material is described in classical control theory textbooks and robotics books as standard form. Thus, some standard theorems, lemmas and corollaries, which are available in references, are sometimes given without a proof. This chapter serves as a short review and as a convenient reference when necessary. In addition, for the robotic control, stability analysis is the key core for all the closed-loop system, therefore, some metric or norms need to be defined such that system could be measured. Those norms that are defined to easily manipulate for the control design, also, all norms have some physical significance.

2.2 Matrix Algebra

Matrix Definition A vector $\mathbf{x}$ defined as a column $(n \times 1)$ vector and $m \times n$ matrix $\mathbf{X}$ that has m rows and n columns are given as follows:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \; x_2 \; \ldots \; x_n \end{bmatrix}^T, \quad \mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{2.1}$$

Matrix $\mathbf{X}$ is defined as a square matrix if $m = n$, that is, $\mathbf{X}$ is an $m \times m$ matrix. Vectors may also be viewed as particular matrix $(n = 1)$, and x_{ij} is an entry from the ith row and jth column of the matrix $\mathbf{X}$. The transpose $\mathbf{X}^T$ of $m \times n$ matrix $\mathbf{X}$ is an $n \times m$ matrix, the entry x_{ij} is the element x_{ji} of the matrix $\mathbf{X}$. Moreover, the related properties are listed as follows:

$$\begin{aligned} \mathbf{X}^T = [x_{ji}], \qquad (\mathbf{X}^T)^T = \mathbf{X}, \\ (\mathbf{XY})^T = \mathbf{Y}^T\mathbf{X}^T, \qquad [\mathbf{X}+\mathbf{Y}]^T = \mathbf{X}^T + \mathbf{Y}^T \end{aligned} \tag{2.2}$$

Z. Li et al., *Advanced Control of Wheeled Inverted Pendulum Systems*,
DOI 10.1007/978-1-4471-2963-9_2, © Springer-Verlag London 2013

Definition 2.1 The square matrix $\mathbf{X} \in \mathbb{R}^{n\times n}$ is said to be

(1) positive semi-definite (denoted by $\mathbf{X} \geq 0$) if $a^T\mathbf{X}a \geq 0$, $\forall a \in \mathbb{R}^n$;
(2) positive definite if $a^T\mathbf{X}a > 0$ for all nonzero $a \in \mathbb{R}^n$;
(3) negative semi-definite if $-\mathbf{X}$ is real positive semi-definite;
(4) negative definite if $-\mathbf{X}$ is positive definite;
(5) indefinite if $\mathbf{X}$ is positive for some $a \in \mathbb{R}^n$ and negative for other $a \in \mathbb{R}^n$;
(6) symmetric if $\mathbf{X}^T = \mathbf{X}$;
(7) skew-symmetric if $\mathbf{X}^T = -\mathbf{X}$;
(8) symmetric positive definite (semi-definite) if $\mathbf{X} > 0$ (≥ 0) and $\mathbf{X} = \mathbf{X}^T$; and
(9) a time-varying matrix $\mathbf{X}(t)$ is uniformly positive definite if there exists $a > 0$ such that $\mathbf{X}(t) \geq a\mathbf{I}$.

The product of $\mathbf{XY}$ of $m \times n$ matrix $\mathbf{X}$ by an $n \times r$ matrix $\mathbf{Y}$ is an $(m \times r)$ matrix $\mathbf{Z}$:

$$\mathbf{Z} = \mathbf{XY}, \quad z_{ij} = \sum_{p=1}^{n} x_{ip} y_{pj} \tag{2.3}$$

and satisfies

$$\mathbf{XY} \neq \mathbf{YX}, \qquad (\mathbf{XYZ}) = \mathbf{X}(\mathbf{YZ}), \qquad (\mathbf{X}+\mathbf{Y})\mathbf{Z} = \mathbf{XZ} + \mathbf{YZ} \tag{2.4}$$

If the matrices $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are symmetric, the following properties are satisfied:

$$\mathbf{XY} = \mathbf{YX}, \qquad \mathbf{XYZ} = \mathbf{YXZ} = \mathbf{ZXY} = \mathbf{ZYX} \tag{2.5}$$

Inner and Outer Product The inner product of two n-dimensional vector $\mathbf{x}$ and $\mathbf{y}$ is a scalar $\mathbf{z} = \mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$. The n-dimensional Euclidean space, denoted by E^n, is $\mathbb{R}^n$ with the inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i \tag{2.6}$$

The outer product of $\mathbf{x}$ and $\mathbf{y}$ is a matrix $\mathbf{Z}$ ($\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$):

$$\mathbf{Z} = \mathbf{xy}^T = \begin{bmatrix} x_1 y_1 & x_2 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix} \tag{2.7}$$

Linear Independent of Vectors The vectors $\mathbf{x}$ are linearly independent if

$$\sum_{i=1}^{n} \alpha_i \mathbf{x} = 0 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0 \tag{2.8}$$

The columns (rows) of $\mathbf{X} \in \mathbb{R}^{m\times n}$ are linearly independent if and only if $\mathbf{X}^T\mathbf{X}$ is a nonsingular matrix, i.e., $\det(\mathbf{X}^T\mathbf{X}) = |\mathbf{X}^T\mathbf{X}| \neq 0$.

Vector Norms Vector norms are positive scalars and are used as measures of length, size, distance, and so on, depending on context. An L_p norm is p-norm of an $(n \times 1)$ vector x defined as

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}, \quad \text{for } 1 \le p < \infty \tag{2.9}$$

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|, \quad \text{for } p = \infty \tag{2.10}$$

The three most commonly used norms are $\|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$ and $\|\mathbf{x}\|_\infty$, which are defined as

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| \tag{2.11}$$

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T\mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} \tag{2.12}$$

$$\|\mathbf{x}\|_\infty = \max |x_i| \tag{2.13}$$

All p-norms are equivalent in the sense that if $\|\cdot\|_{p_1}$ and $\|\cdot\|_{p_2}$ are two different p-norms, then there exist positive constants c_1 and c_2 such that

$$c_1 \|\mathbf{x}\|_{p_1} \le \|\mathbf{x}\|_{p_2} \le c_2 \|\mathbf{x}\|_{p_1}, \quad \forall \mathbf{x} \in \mathbb{R}^n \tag{2.14}$$

Determinants of Matrix The determinant of an $n \times n$ square matrix $\mathbf{X}$ is defined as the sum of the signed products of all possible combinations of n elements, in which each element is taken from a different row and column. The determinant of $\mathbf{X}$ is denoted by $\det(\mathbf{X})$ as

$$\det(\mathbf{X}) = \sum^{n!} (-1)^p x_{1p_1} x_{1p_2} \ldots x_{1p_n} \tag{2.15}$$

where $p_1, p_2, \ldots, p_n$ is a permutation of $1, 2, \ldots, n$ and the sum is taken over all possible permutations. A permutation is a rearrangement of $1, 2, \ldots, n$ into some other order, such as $n, 1, \ldots, 2$, that is obtained by successive transpositions. A transposition is the interchange of places of two numbers in the list $1, 2, \ldots, n$. The exponent p of -1 is the number of transpositions it takes to go from the natural order to $p_1, p_2, \ldots, p_n$. There are $n!$ possible permutations of n numbers, so each determinant is the sum of $n!$ products.

Eigenvalues and Properties of Matrix For a given general matrix $\mathbf{X} \in \mathbb{R}^{n \times n}$, there can be up to n such special vectors satisfying $\det(\mathbf{X} - \lambda \mathbf{I}) = 0$. These vectors $\lambda_1, \ldots, \lambda_i, \ldots, \lambda_n$ are called the eigenvectors of $\mathbf{X}$, and the proportionality constants are called the eigenvalues.

The properties for the eigenvectors can be obtained as

- If $\lambda_i(\mathbf{X}) > 0$ (≥ 0), then $\mathbf{X} = \mathbf{X}^T \in \mathbb{R}^{n \times n}$ is positive (semi-)definite.
- If $\lambda_i(\mathbf{X}) < 0$ (≤ 0), $\mathbf{X} = \mathbf{X}^T \in \mathbb{R}^{n \times n}$ is negative (semi-)definite.

- If a matrix is symmetric, then its eigenvalues are all real.
- A necessary condition for a square matrix $\mathbf{X}$ to be positive definite is that its diagonal elements be strictly positive.
- A necessary and sufficient condition for a symmetric matrix $\mathbf{X}$ to be positive definite is that all its principal minors be strictly positive.
- If $\mathbf{X} = \mathbf{X}^T > 0$ (≥ 0) and $\mathbf{Y} = \mathbf{Y}^T > 0$ (≥ 0), then $\mathbf{X} + \mathbf{Y} > 0$ (≥ 0) and has all eigenvalues real positive but it is not true in general that $\mathbf{XY} > 0$ (≥ 0). If and only if $\mathbf{X}$ and $\mathbf{Y}$ are commutative, i.e., $\mathbf{XY} = \mathbf{YX}$, then $\mathbf{XY} > 0$ (≥ 0).
- If matrix $\mathbf{X}$ is symmetric positive semi-definite, then it can be decomposed as $\mathbf{X} = U^T \Lambda U$, where U is a unitary matrix and satisfies $U^T U = I$, and Λ is a diagonal matrix containing the eigenvalues of the matrix $\mathbf{X}$, moreover, the following inequality is satisfied $\lambda_{\min}(\mathbf{X})\|a\|^2 \leq a^T \mathbf{X} a \leq \lambda_{\max}(\mathbf{X})\|a\|^2$.
- If $\mathbf{X}$ is skew-symmetric, then $a^T \mathbf{X} a = 0$ for all $a \in \mathbb{R}^n$.
- If $\mathbf{X} \geq 0$, then $\mathbf{X} + \mathbf{X}^T$ is symmetric positive semi-definite.
- If matrix $\mathbf{X}$ is square, then it can be represented as the sum of a symmetric matrix and a skew-symmetric matrix as follows: $\mathbf{X} = \frac{\mathbf{X}+\mathbf{X}^T}{2} + \frac{\mathbf{X}-\mathbf{X}^T}{2}$.
- Since $a^T \mathbf{X} a = a^T \frac{\mathbf{X}+\mathbf{X}^T}{2} a$, the positive definiteness of a matrix can be determined by only the symmetric part of $\mathbf{X}$.
- If $\det(\mathbf{X}) \neq 0$, the square matrix $\mathbf{X}$ is non-singular. Otherwise, $\mathbf{X}$ is singular.

Diagonal Matrix Given $D = \mathrm{diag}[d_{ii}] \in \mathbb{R}^{n \times n}$, $d_{ii} > 0$, $i = 1, \ldots, n$, we have the following inequality

$$0 < \gamma_1 x^T x \leq x^T D x \leq \gamma_2 x^T x, \quad \forall x \in \mathbb{R}^n \tag{2.16}$$

where both γ_1 and γ_2 are positive constants, and $\gamma_1 \leq \min_{i=1,\ldots,n}\{d_{ij}\}$ and $\gamma_2 \geq \max_{i=1,\ldots,n}\{d_{ii}\}$.

Lemma 2.2 *For* $K = \mathrm{diag}[k_{ii}] \in \mathbb{R}^{n \times n}$, *and* $a = [a_1, a_2, \ldots, a_n]^T \in \mathbb{R}^n$, *if* $u = -K\,\mathrm{sgn}(x)$, *and* $k_{ii} \geq |d_i|$, *then* $x^T D(u - a) \leq 0$.

Lemma 2.3 *For* $K = \mathrm{diag}[k_{ii}] \in \mathbb{R}^{n \times n}$ *and* $\underline{k} = \min_{i=1,\ldots,n}\{k_{ii}\}$, *if* $u = -K\,\mathrm{sgn}_1(x)$, *and* $\underline{k} \geq \frac{\|D\|\|a\|}{\gamma_1}$ (*or* $\geq \frac{\gamma_2}{\gamma_1}\|a\|$), *then* $x^T D(u - a) \leq 0$.

Symmetric Positive Definite Matrix

Lemma 2.4 *Consider* $u = -k\,\mathrm{sgn}_1(x)$, *and* $0 < \gamma_1 x^T x \leq x^T D x \leq \gamma_2 x^T x$, *if* $k \geq \frac{\|D\|\|a\|}{\gamma_1}$ (*or* $\geq \frac{\gamma_2}{\gamma_1\|a\|}$), *then* $x^T D(u - a) \leq 0$.

Although $D(t)$ is time variant and usually unknown, the lower and upper bounds of D can be obtained from the view point of physical model for practical systems, therefore, it can be chosen such that $\gamma_1 \leq \lambda_{\min}(D)$ and $\gamma_2 \geq \lambda_{\max}(D)$. If D is symmetric negative definite, we have

$$-\gamma_2 x^T x \leq x^T D x \leq -\gamma_1 x^T x < 0, \quad \forall x \in \mathbb{R}^n \tag{2.17}$$

where γ_1 and γ_2 are positive constants.

Induced Norm of Matrices For an $m \times n$ matrix $\mathbf{X}$, the induced p-norm of $\mathbf{X}$ is defined by

$$\|\mathbf{X}\|_p = \sup_{x \neq 0} \frac{\|\mathbf{X}y\|_p}{\|y\|_p} = \sup_{\|y\|_p = 1} \|\mathbf{X}y\|_p \tag{2.18}$$

As similar as the vector norms, for $p = 1, 2, \infty$, the corresponding induced norms could be obtained as follows:

$$\|\mathbf{X}\|_1 = \max_j \sum_{i=1}^{m} |x_{ij}| \quad \text{(column sum)} \tag{2.19}$$

$$\|\mathbf{X}\|_2 = \max_i \sqrt{\lambda_i \left(\mathbf{X}^T \mathbf{X}\right)} \tag{2.20}$$

$$\|\mathbf{X}\|_\infty = \max_i \sum_{j=1}^{m} |x_{ij}| \quad \text{(row sum)} \tag{2.21}$$

The induced norms are also satisfying the following:

$$\|\mathbf{XY}\|_p \leq \|\mathbf{X}\|_p \|\mathbf{Y}\|_p, \quad \forall \mathbf{X} \in \mathbb{R}^{n \times m}, \mathbf{Y} \in \mathbb{R}^{m \times l} \tag{2.22}$$

2.3 Norms for Functions

Open and Closed Sets

Definition 2.5 A set $S \subset \mathbb{R}^n$ is called open, if for each $x \in S$ there exists and $\epsilon > 0$ such that the interval $(x - \epsilon, x + \epsilon)$ is contained in S. Such an interval is often called an ϵ-neighborhood of x, or simply a neighborhood of x.

Continuous Function

Definition 2.6 A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be continuous at a point $x \in \mathbb{R}^n \subseteq \mathbb{R}^m$, if for each $\epsilon > 0$, there exists a $\delta(\epsilon, x)$ such that for all $y \in \mathbb{R}^n$ satisfying $\|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon$. A function f is continuous on $\mathbb{R}^n$ if it is continuous at every point in $\mathbb{R}^n$.

Definition 2.7 A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is uniformly continuous on $\mathbb{R}^n \subseteq \mathbb{R}^m$ if for each ϵ, there exists a $\delta(\epsilon)$ such that for all $x, y \in \mathbb{R}^n$ satisfying $|x - y| < \delta(\epsilon)$, then $|f(x) - f(y)| < \epsilon$.

Definition 2.8 A function $f : [0, \infty) \to R$ is piecewise continuous on $[0, \infty)$ if f is continuous on any finite interval $[a, b] \subset [0, \infty)$ except at a finite number of points on each of these intervals.

Definition 2.9 A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be Lipschitz continuous if there exists a constant $L > 0$, which is sometimes called the Lipschitz constant, such that $|f(x) - f(y)| \le L|x - y|$ for all $x, y \in \mathbb{R}^n$, where $\mathbb{R}^n \subseteq \mathbb{R}^m$.

A function f is continuous on a set of S if it is continuous at every point of S, and it is uniformly continuous on S if given $\epsilon > 0$, there is $\delta(\epsilon) > 0$ (dependent only on ϵ), such that the inequality holds for all $x, y \in S$.

Differentiable Function The derivative of function $f : \mathbb{R} \to \mathbb{R}$ at a point x is defined

$$\dot{f}(x) = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x} \tag{2.23}$$

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable at point x (on a set S) if the partial derivatives $\partial f_i / \partial x_j$ exist and continuous at x (at every point of S) for $1 \le i \le m$, $1 \le j \le n$, and the Jacobian matrix $f : \mathbb{R}^n \to \mathbb{R}^m$ is defined as

$$J = \left[\frac{\partial f}{\partial x}\right] = \begin{bmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial f_m/\partial x_1 & \cdots & \partial f_m/\partial x_n \end{bmatrix} \in \mathbb{R}^{m \times n} \tag{2.24}$$

For a scalar function $f(x, y)$ that depends on x and y, the gradient with respect to x is defined as

$$\nabla_x f(x, y) = \left[\frac{\partial f}{\partial x}\right] = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right] \tag{2.25}$$

If f is only a function of x, we denote $\nabla f(x)$.

Mean Value Theorem If a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at each point $x, y \in \mathbb{R}^n$, where $\mathbb{R}^n \subseteq \mathbb{R}^m$, such that the line segment $L(x, y) \in \mathbb{R}^n$, then there exists some $z \in L(x, y)$ such that

$$f(y) - f(x) = \left.\frac{\partial f(x)}{\partial x}\right|_{x=z} (y - x) \tag{2.26}$$

which is equivalent to $\frac{\partial f(x)}{\partial x} = \frac{f(y)-f(x)}{y-x}$.

Function Norms Given the time varying function $x(t) : \mathbb{R}_+ \to \mathbb{R}$, be continuous or piecewise continuous. The p-norm for $x(t)$ is defined as

$$\left\|x(t)\right\|_p = \left(\int_0^\infty \left|x(t)^p\right| dt\right)^{1/p}, \quad \text{for } p \in [1, \infty) \tag{2.27}$$

$$\left\|x(t)\right\|_\infty = \sup\nolimits_{t \in [0,\infty)} \left|x(t)\right|, \quad \text{for } p = \infty \tag{2.28}$$

The function space over which the signal norm exists is define by letting $p = 1, 2, \infty$, the corresponding norm spaces are called L_1, L_2, L_∞, respectively, let $x(t) \in [0, \infty)$, which is define by

$$L_p = \left\{x(t) \in \mathbb{R}^n : \|x\|_p < \infty\right\} \tag{2.29}$$

for $p \in [1, \infty)$, that is, L_p is the set of all vector functions in $\mathbb{R}^n$ for which the p-norm is well defined (finite). From a signal point of view, the 1-norm, $\|x\|_1$, of the signal $x(t)$ is the integral of its absolute value, the square $\|x\|_2^2$ of the 2-norm is often called the energy of the signal $x(t)$, and the ∞-norm is its absolute maximum amplitude or peak value.

The following inequalities for signals will be useful:

- *Hölder's Inequality*: If scalar time varying function $x \in L_p$ and $y \in L_p$ for $p, q \in [1, +\infty)$ and $1/p + 1/q = 1$, then $xy \in L_1$ and $\|xy\|_1 \leq \|x\|_p \|y\|_q$.
- *Minkowski Inequality*: If scalar time varying function, $x, y \in L_p$ for $p \in [1, +\infty)$, then $x + y \in L_p$ and $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.
- *Young's Inequality*: For scalar time varying function $x(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$, it holds that $2xy \leq \frac{1}{\epsilon}x^2 + \epsilon y^2$, for any $\epsilon > 0$.
- *Completing the Square*: For scalar time functions $x(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$, we have $-x^2 + 2xy = -x^2 + 2xy - y^2 + y^2 \leq y^2$.

Let $f(t) : \mathbb{R}_+ \to \mathbb{R}^n$, $f = [f_1(t), f_2(t), \ldots, f_n(t)]^T$ be a continuous or piecewise continuous vector function. Then, the corresponding p-norm spaces are defined as

$$L_p^n \triangleq \left\{f : \mathbb{R}_+ \to \mathbb{R}^n \;\middle|\; \|f\|_p = \int_0^\infty \|f\|^p \, dt < \infty, \text{ for } p \in [1, \infty)\right\} \tag{2.30}$$

$$L_\infty^n \triangleq \left\{f : \mathbb{R}_+ \to \mathbb{R}^n \;\middle|\; \|f\|_\infty = \sup_{t \in [0,\infty)} \|f\|_\infty < \infty\right\} \tag{2.31}$$

In some cases, during the period $[0, T]$, signals are bounded over finite time intervals, but may become infinity as time goes to infinity. Therefore, we only choose this period to extend L_p^n spaces by

$$L_b^n \triangleq \left\{f : \mathbb{R}_+ \to \mathbb{R}^n \;\middle|\; \|f\|_{pT} = \int_0^T \|f\|^p \, dt < \infty, \ \forall T \in \mathbb{R}_+\right\} \tag{2.32}$$

$$L_{b\infty}^n \triangleq \left\{f : \mathbb{R}_+ \to \mathbb{R}^n \;\middle|\; \|f\|_\infty = \sup_{t \leq T} \|f\| < \infty, \ \forall T \in \mathbb{R}_+\right\} \tag{2.33}$$

Definition and Properties of Sign Functions The signum function of a real number x is defined as follows:

$$\operatorname{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases} \tag{2.34}$$

The properties of sign function are listed as follows.

(1) Any real number can be expressed as the product of its absolute value and its sign function:

$$x = \operatorname{sgn}(x) \cdot |x| \tag{2.35}$$

whenever x is not equal to 0, from Eq. (2.35), we have

$$\mathrm{sgn}(x) = \frac{x}{|x|} \tag{2.36}$$

(2) The signum function is the derivative of the absolute value function (up to the indeterminacy at zero): Note, the resultant power of x is 0, similar to the ordinary derivative of x. The numbers cancel and all we are left with is the sign of x,

$$\frac{d|x|}{dx} = \mathrm{sgn}(x) \tag{2.37}$$

(3) The signum function is differentiable with derivative 0 everywhere except at 0. It is not differentiable at 0 in the ordinary sense, but under the generalized notion of differentiation in distribution theory, the derivative of the signum function is two times δ function,

$$\frac{d\,\mathrm{sgn}(x)}{dx} = 2\delta \tag{2.38}$$

where

$$\delta(x) = \begin{cases} +\infty, & x = 0 \\ 0, & x \neq 0, \end{cases} \qquad \int_{-\infty}^{+\infty} \delta(x)\,dx = 1 \tag{2.39}$$

(4) For $k \geq 0$, a smooth approximation of the sign function is

$$\mathrm{sgn}(x) \approx \tanh(kx) \tag{2.40}$$

(5) For vectors, the sign function is defined as

$$\mathrm{sgn}(\mathbf{X}) \triangleq \left[\mathrm{sgn}(x_1)\ \ \mathrm{sgn}(x_2)\ \ \ldots\ \ \mathrm{sgn}(x_n)\right]^T \tag{2.41}$$

(6) In general, $\mathrm{sgn}(\mathbf{x}) \neq \mathrm{sgn}_1(x)$. However, $\mathrm{sgn}(\mathbf{x}) = \mathrm{sgn}(x)$ and $x^T\,\mathrm{sgn}(x) = x^T\,\mathrm{sgn}_1(x)$ if and only if one element of x is nonzero and the remaining elements of x are zero.

(7) The L_2-norm of $\mathrm{sgn}(x)$ is 1 and $\|\,\mathrm{sgn}(x)\| = \sqrt{n}$. From $x^T\,\mathrm{sgn}(x) = \sum_{j=1}^{n} |x_j|$ $= \|x\|_1$ and $x^T\,\mathrm{sgn}(x) = \frac{\|x\|^2}{\|x\|} = \|x\|$.

(8) While $\mathrm{sgn}(x)$ defines a unit vector in the direction of x, $\mathrm{sgn}(x)$ maps all the vectors that are in the same quadrant (excluding those vectors on the axes) into one vector, $\mathrm{sgn}(x)$ covers an n-dimensional unit ball. In other words, $\mathrm{sgn}(x)$ defines $3^n - 1$ vectors evenly distributed in the n-dimensional space.

(9) If $M \in \mathbb{R}^{n\times n}$ is positive definite (negative definite), i.e., $x^T Mx \geq 0$ (≤ 0), then $x^T M\,\mathrm{sgn}_1(x)$ is also positive (negative) and well defined because $x^T M\,\mathrm{sgn}_1(x)$ $= \frac{x^T Mx}{\|x\|}$ while no conclusion can be drawn for $x^T M\,\mathrm{sgn}(x)$.

2.4 Definitions

Consider the following nonautonomous system:

$$\dot{x} = f(t, x) \tag{2.42}$$

where $f : [0, \infty) \times D \to \mathbb{R}^n$ is piecewise continuous in t and locally Lipschitz in x on $[0, \infty) \times D$ and $D \subseteq \mathbb{R}^n$ is a domain that contains the origin $x = 0$.

Definition 2.10 The origin $x = 0$ is the equilibrium point of (2.42) if

$$f(t, 0) = 0, \quad \forall t \geq 0 \tag{2.43}$$

Definition 2.11 A continuous function $\alpha : [0, a) \to \mathbb{R}^+$ is said to belong to class K if it is strictly increasing and $\alpha(0) = 0$. It is said to belong to class K_∞ if $a = \infty$ and $\alpha(r) \to \infty$ as $r \to \infty$.

Definition 2.12 A continuous function $\beta : [0, a) \times \mathbb{R}^+ \to \mathbb{R}^+$ is said to belong to class KL if, for each fixed s, the mapping $\beta(r, s)$ belongs to class K with respect to r and, for each fixed r, the mapping $\beta(r, s)$ is decreasing with respect to s and $\beta(r, s) \to 0$ as $s \to \infty$. It is said to belong to class KL_∞ if, in addition, for each fixed s the mapping $\beta(r, s)$ belongs to class K_∞ with respect to r.

Definition 2.13 The equilibrium point $x = 0$ of (2.42) is

(1) stable if, for each $\varepsilon > 0$, there is $\delta = \delta(\varepsilon, t_0) > 0$ such that

$$\|x(t_0)\| < \delta \quad \Rightarrow \quad \|x(t)\| < \varepsilon, \quad \forall t \geq t_0 \geq 0 \tag{2.44}$$

(2) uniformly stable if, for each $\varepsilon > 0$, there is $\delta = \delta(\varepsilon) > 0$ independent of t_0 such that (2.44) is satisfied;
(3) unstable if it is not stable;
(4) asymptotically stable if it is stable and there is a positive constant $c = c(t_0)$ such that $x(t) \to 0$ as $t \to \infty$, for all $\|x(t_0)\| < c$;
(5) uniformly asymptotically stable if it is uniformly stable and there is a positive constant c, independent of t_0, such that for all $\|x(t_0)\| < c$, $x(t) \to 0$ as $t \to \infty$, uniformly in t_0; that is, for each $\eta > 0$, there is $T = T(\eta) > 0$ such that

$$\|x(t)\| < \eta, \quad \forall t \geq t_0 + T(\eta), \ \forall \|x(t_0)\| < c \tag{2.45}$$

(6) globally uniformly asymptotically stable (GUAS) if it is uniformly stable, $\delta(\varepsilon)$ can be chosen to satisfy $\lim_{\varepsilon \to \infty} \delta(\varepsilon) = \infty$, and, for each pair of positive numbers η and c, there is $T = T(\eta, c) > 0$ such that

$$\|x(t)\| < \eta, \quad \forall t \geq t_0 + T(\eta, c), \ \forall \|x(t_0)\| < c \tag{2.46}$$

Definition 2.14 The equilibrium point $x = 0$ of (2.42) is exponentially stable if there exist positive constants c, k, and λ such that

$$\|x(t)\| \leq k\big(\|x(t_0)\|\big)e^{-\lambda(t-t_0)}, \quad \forall t \geq t_0 \geq 0, \ \forall \|x(t_0)\| < c \tag{2.47}$$

and GES if (2.47) is satisfied for any initial state $x(t_0)$.

Definition 2.15 The equilibrium point $x = 0$ of (2.42) is K-exponentially stable if there exist positive constants c and λ and a class K function α such that

$$\|x(t)\| \leq \alpha\big(\|x(t_0)\|\big)e^{-\lambda(t-t_0)}, \quad \forall t \geq t_0 \geq 0, \ \forall \|x(t_0)\| < c \tag{2.48}$$

and globally K-exponentially stable if (2.48) is satisfied for any initial state $x(t_0)$.

Definition 2.16 The solutions of (2.42) are as follows:

(1) uniformly bounded if there exists a positive constant c, independent of $t_0 \geq 0$, and for every $a \in (0, c)$, there is $\beta = \beta(a) > 0$, independent of t_0, such that

$$\|x(t_0)\| \leq a \quad \Rightarrow \quad \|x(t)\| \leq \beta, \quad \forall t \geq t_0 \tag{2.49}$$

(2) globally uniformly bounded if (2.49) holds for an arbitrarily large a;
(3) uniformly ultimately bounded with ultimate bound b if there exist positive constants b and c, independent of $t_0 \geq 0$, and for every $a \in (0, c)$ there is $T = T(a, b) \geq 0$, independent of t_0, such that

$$\|x(t_0)\| \leq a \quad \Rightarrow \quad \|x(t)\| \leq b, \quad \forall t \geq t_0 + T \tag{2.50}$$

(4) globally uniformly ultimately bounded if (2.50) holds for an arbitrarily large a.

2.5 Lemmas and Theorems

Lemma 2.17 *The equilibrium point $x = 0$ of (2.42) is*

(1) *uniformly stable if and only if there exist a class K function α and a positive constant c, independent of t_0, such that*

$$\|x(t)\| \leq \alpha\big(\|x(t_0)\|\big), \quad \forall t \geq t_0 \geq 0, \ \forall x(t_0) \mid \|x(t_0)\| < c \tag{2.51}$$

(2) *uniformly asymptotically stable if and only if there exist a β function and a positive constant c, independent of t_0, such that*

$$\|x(t)\| \leq \beta\big(\|x(t_0)\|, t - t_0\big), \quad \forall t \geq t_0 \geq 0, \ \forall x(t_0) \mid \|x(t_0)\| < c \tag{2.52}$$

(3) *GUAS if and only if inequality (2.52) is satisfied with $\beta \in KL_\infty$ for any initial state $x(t_0)$.*

Lemma 2.18 *Assume that $d : \mathbb{R}^n \to \mathbb{R}^n$ satisfies*

$$P\left[\frac{\partial d}{\partial x}\right] + \left[\frac{\partial d}{\partial x}\right]^T P \geq 0, \quad \forall x \in \mathbb{R}^n \tag{2.53}$$

when $P = P^T > 0$. Then

$$(x - y)^T P\big(d(x) - d(y)\big) \geq 0, \quad \forall x, y \in \mathbb{R}^n \tag{2.54}$$

Theorem 2.19 *Let $D = \{x \in \mathbb{R}^n \mid \|x\| < r\}$ and $x = 0$ be an equilibrium point of (2.42). Let $V : D \times \mathbb{R}^n \to \mathbb{R}^+$ be a continuously differentiable function such that $\forall t \geq 0, \forall x \in D$,*

$$\begin{aligned} \gamma_1\big(\|x\|\big) \leq V(x,t) \leq \gamma_2\big(\|x\|\big) \\ \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(t,x) \leq -\gamma_3\big(\|x\|\big) \end{aligned} \tag{2.55}$$

Then the equilibrium point $x = 0$ is

(1) *uniformly stable, if γ_1 and γ_2 are class K functions on $[0, r)$ and $\gamma_3 \geq 0$ on $[0, r)$;*
(2) *uniformly asymptotically stable, if γ_1, γ_2 and γ_3 are class K functions on $[0, r)$;*
(3) *exponentially stable if $\gamma_i(\rho) = k_i \rho^\alpha$ on $[0, r)$, $k_i > 0$, $\alpha > 0$, $i = 1, 2, 3$;*
(4) *globally uniformly stable if $D = \mathbb{R}^n$, γ_1 and γ_2 are class K_∞ functions, and $\gamma_3 \geq 0$ on $\mathbb{R}^+$;*
(5) *GUAS if $D = \mathbb{R}^n$, γ_1 and γ_2 are class K_∞ functions, and γ_3 is a class K function on $\mathbb{R}^+$;*
(6) *GES, if $D = \mathbb{R}^n$, $\gamma_i(\rho) = k_i \rho^\alpha$ on $\mathbb{R}^+$, $k_i > 0$, $\alpha > 0$, $i = 1, 2, 3$.*

Theorem 2.20 *Let $x = 0$ be an equilibrium point of (2.42) and suppose that f is locally Lipschitz in x and uniformly continuous in t. Let $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^+$ be a continuously differentiable function such that*

$$\begin{aligned} \gamma_1\big(\|x\|\big) \leq V(x,t) \leq \gamma_2\big(\|x\|\big) \\ \dot{V} = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(t,x) \leq -W(x) \leq 0 \end{aligned} \tag{2.56}$$

for all $t \geq 0$ and $x \in \mathbb{R}^n$, where γ_1 and γ_2 are class K_∞ functions, and W is a continuous function. Then all solutions of (2.42) are globally uniformly bounded and satisfy

$$\lim_{t \to \infty} W\big(x(t)\big) = 0 \tag{2.57}$$

In addition, if $W(x)$ is positive definite, then the equilibrium point $x = 0$ is GUAS.

Theorem 2.21 *Let $V : [0, \infty) \times D \to \mathbb{R}$ be a continuously differentiable function and $D \in \mathbb{R}^n$ be a domain that contains the origin such that*

$$\begin{aligned} \alpha_1\big(\|x\|\big) \leq V(x,t) \leq \alpha_2\big(\|x\|\big) \\ \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(t,x) \leq -W(x), \quad \forall \|x\| \geq \mu > 0 \end{aligned} \tag{2.58}$$

for all $t \geq 0$ and $x \in D$ where α_1 and α_2 are class K functions, and W is a continuously positive definite function. Take $r > 0$ such that $B_r \subset D$ and suppose that

$$\mu < \alpha_2^{-1}\big(\alpha_1(r)\big) \tag{2.59}$$

Then, there exists a class KL function β and for every initial state $x(t_0)$, satisfying $\|x(t_0)\| < \alpha_2^{-1}(\alpha_1(r))$, there is $T > 0$ (dependent on $x(t_0)$ and μ) such that the solutions of (2.42) satisfies

$$\begin{aligned} \|x(t)\| &\le \beta\big(\|x(t_o)\|, t - t_0\big), \quad \forall t_0 \le t \le t_0 + T \\ \|x(t)\| &< \alpha_1^{-1}\big(\alpha_2(\mu)\big), \quad \forall t \ge t_0 + T \end{aligned} \tag{2.60}$$

Moreover, if $D = \mathbb{R}^n$ and α_1 belongs to class K_∞, then (2.60) holds for any initial state $x(t_0)$ with no restriction on how large μ is.

2.6 Input-to-State Stability

Definition 2.22 The system

$$\dot{x} = f(t, x, u) \tag{2.61}$$

where f is piecewise continuous in t and locally Lipschitz in x and u, is said to be input-to-state stable (ISS) if there exist a class KL function β and a class K function γ, such that, for any $x(t_0)$ and for any input $u(\cdot)$ continuous and bounded on $[0, \infty)$, the solution exists for all $t \ge t_0 \ge 0$ and satisfies

$$\|x(t)\| \le \beta\big(\|x(t_0)\|, t - t_0\big) + \gamma\Big(\sup_{t_0 \le \tau \le t} \|u(\tau)\|\Big) \tag{2.62}$$

The following theorem establishes the equivalence between the existence of a Lyapunov-like function and the input-to-state stability.

Theorem 2.23 *Suppose that for the system (2.61) there exists a C^1 function $V : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^+$ such that for all $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$,*

$$\begin{aligned} \gamma_1\big(\|x\|\big) \le V(t, x) &\le \gamma_2\big(\|x\|\big) \\ \|x\| \ge \rho\big(\|u\|\big) \quad \Rightarrow \quad \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(t, x, u) &\le -\gamma_3\big(\|x\|\big) \end{aligned} \tag{2.63}$$

where γ_1, γ_2 and ρ are class K_∞ functions and γ_3 is a class-K function. Then the system (2.61) is ISS with $\gamma = \gamma_1^{-1} \circ \gamma_2 \circ \rho$.

Proof If $x(t_0)$ is in the set

$$R_{t_0} = \Big\{x \in \mathbb{R}^n \;\Big|\; \|x\| \le \rho\big(\sup\|u(\tau)\|_{\tau \ge \tau_0}\big)\Big\} \tag{2.64}$$

then $x(t)$ remains within the set

$$S_{t_0} = \Big\{x \in \mathbb{R}^n \;\Big|\; \|x\| \le \gamma_1^{-1} \circ \gamma_2 \circ \rho\big(\sup\|u(\tau)\|_{\tau \ge \tau_0}\big)\Big\} \tag{2.65}$$

for all $t \ge t_0$. Define $B = [t_0, T)$ as the time interval before $x(t)$ enters R_{t_0} for the first time. In view of the definition of R_{t_0}, we have

$$\dot{V} \le -\gamma_3 \circ \gamma_2^{-1}(V), \quad \forall t \in B \tag{2.66}$$

Then, there exists a class-KL function β_v such that $V(t) \le \beta_v(V(t_0), t - t_0)$, $\forall t \in B$, which implies

$$\|x(t)\| \le \gamma_1^{-1}\big(\beta_v\big(\gamma_2\big(\|x(t_0)\|\big), t - t_0\big)\big) := \beta\big(\|x(t_0)\|, t - t_0\big), \quad \forall t \in B \tag{2.67}$$

On the other hand, by (2.65), we conclude that

$$\|x(t)\| \le \gamma_1^{-1} \circ \gamma_2 \circ \rho\big(\sup\|u(\tau)\|_{\tau \ge \tau_0}\big) := \gamma\big(\sup\|u(\tau)\|_{\tau \ge \tau_0}\big) \tag{2.68}$$

for all $t \in [t_0, \infty]$. Then by (2.67) and (2.68),

$$\|x(t)\| \le \beta\big(\|x(t_0)\|, t - t_0\big) + \gamma\big(\sup\|u(\tau)\|_{\tau \ge \tau_0}\big), \quad \forall t \ge t_0 \ge 0 \tag{2.69}$$

By causality, we have

$$\|x(t)\| \le \beta\big(\|x(t_0)\|, t - t_0\big) + \gamma\big(\sup\|u(\tau)\|_{t_0 \le \tau \le t}\big), \quad \forall t \ge t_0 \ge 0 \tag{2.70}$$

A function V satisfying conditions (2.63) is called an ISS Lyapunov function. □

2.7 Lyapunov's Direct Method

This section presents an extension of the Lyapunov function concept, which is a useful tool to design an adaptive controller for nonlinear systems. Assuming that the problem is to design a feedback control law $\alpha(x)$ for the time-invariant system:

$$\dot{x} = f(x, u), \quad x \in \mathbb{R}^n, \ u \in \mathbb{R} \tag{2.71}$$

$$f(0, 0) = 0 \tag{2.72}$$

such that the equilibrium $x = 0$ of the closed loop system:

$$\dot{x} = f\big(x, \alpha(x)\big) \tag{2.73}$$

is globally asymptotically stable (GAS). We can take a function $V(x)$ as a Lyapunov candidate function, and require that its derivative along the solutions of (2.73) satisfy $\dot{V}(x) \le -W(x)$, where $W(x)$ is a positive definite function. We therefore need to find $\alpha(x)$ guarantee that for all $x \in \mathbb{R}^n$ such that

$$\frac{\partial V(x)}{\partial x} f\big(x, \alpha(x)\big) \le -W(x) \tag{2.74}$$

This is a difficult problem. A stabilizing control law for (2.72) may exist but we may fail to satisfy (2.74) because of a poor choice of $V(x)$ and $W(x)$. A system for which a good choice of $V(x)$ and $W(x)$ exists is said to possess a control Lyapunov function (CLF). For systems affine in the control:

$$\dot{x} = f(x) + g(x)u, \quad f(0) = 0 \tag{2.75}$$

the CLF inequality (2.74) becomes

$$\frac{\partial V}{\partial x} f(x) + \frac{\partial V}{\partial x} g(x)\alpha(x) \leq -W(x) \tag{2.76}$$

If $V(x)$ is a CLF for (2.75), then a particular stabilizing control law $\alpha(x)$, smooth for all $x \neq 0$, is given by

$$u = \alpha(x) = \begin{cases} -\frac{\frac{\partial V}{\partial x} f(x) + \sqrt{(\frac{\partial V}{\partial x} f(x))^2 + (\frac{\partial V}{\partial x} g(x))^4}}{\frac{\partial V}{\partial x} g(x)}, & \frac{\partial V}{\partial x} g(x) \neq 0 \\ 0, & \frac{\partial V}{\partial x} g(x) = 0 \end{cases}$$

It should be noted that (2.76) can be satisfied only if

$$\frac{\partial V}{\partial x} g(x) = 0 \quad \Rightarrow \quad \frac{\partial V}{\partial x} f(x) < 0, \quad \forall x \neq 0 \tag{2.77}$$

and that in this case (2.77) gives

$$W(x) = \sqrt{\left(\frac{\partial V}{\partial x} f\right)^2 + \left(\frac{\partial V}{\partial x} g\right)^4} > 0, \quad \forall x \neq 0 \tag{2.78}$$

The main drawback of the CLF concept as a design tool is that for most nonlinear systems a CLF is not known. The task of finding an appropriate CLF may be as complex as that of designing a stabilizing feedback law.

2.8 Barbalat-Like Lemmas

This section presents lemmas that are useful in investigating the convergence of time-varying systems.

If a function $f \in L_1$ may not be bounded. On the converse, if a function f is bounded, it is not necessary that $f \in L_1$. However, if $f \in L_1 \cap L_\infty$, then $f \in L_p$ for all $p \in [1, \infty)$. Moreover, $f \in L_p$ could not lead to $f \to 0$ as $t \to \infty$. If f is bounded can also lead to $f \to 0$ as $t \to \infty$. However, we have the following results.

Lemma 2.24 (Barbalat's Lemma) *Consider the function $\phi : \mathbb{R}^+ \to \mathbb{R}$. If ϕ is uniformly continuous and $\lim_{t\to\infty} \int_0^t \phi(\tau)\, d\tau$ exists and is finite, then*

$$\lim_{t\to\infty} \phi(t) = 0 \tag{2.79}$$

Lemma 2.25 *Assume that a nonnegative scalar differentiable function $f(t)$ enjoys the following conditions*

$$\begin{aligned} &1.\ \left|\frac{d}{dt} f(t)\right| \leq k_1 f(t) \\ &2.\ \int_0^\infty f(t)\, dt \leq k_2 \end{aligned} \tag{2.80}$$

for all $t \geq 0$, where k_1 and k_2 are positive constants, then $\lim_{t\to\infty} f(t) = 0$.

Proof Integrating both sides of (2.80) gives

$$\begin{aligned} f(t) &\le f(0) + k_1 \int_0^t f(s)\,ds \le f(0) + k_1 k_2 \\ f(t) &\ge f(0) - k_1 \int_0^t f(s)\,ds \ge f(0) - k_1 k_2 \end{aligned} \tag{2.81}$$

These inequalities imply that $f(t)$ is a uniform bounded function. From (2.81) and the second condition in (2.80), we have that $f(t)$ is also bounded on the half axis $[0, \infty)$, i.e., $f(t) \le k_3$ with k_3 a positive constant. Hence, $|\frac{d}{dt} f(t)| \le k_1 k_3$. Now assume that $\lim_{t\to\infty} f(t) \ne 0$. Then there exists a sequence of points t_i and a positive constant ϵ such that $f(t_i) \ge \epsilon$, $t_i \to \infty$, $i \to \infty$, $|t_i - t_{i-1}| > 2\epsilon/(k_1 k_3)$ and moreover $f(s) \ge \epsilon/2$, $s \in L_i = [t_i - \epsilon/(2k_1k_3), t_i + \epsilon/(2k_1k_3)]$. Since the segments L_i and L_j do not intersect for any i and j with $i \ne j$, we have

$$\int_0^\infty f(t)\,dt \ge \int_0^T f(t)\,dt \ge \sum_{t_i \le T} \int_{L_i} f(t)\,dt \ge \frac{\epsilon}{2} \frac{\epsilon}{k_1 k_3} M(T) \tag{2.82}$$

where $M(T)$ is the number of points t_i not exceeding T. Since $\lim_{T\to\infty} M(T) = \infty$, the integral $\int_0^\infty f(t)\,dt$ is divergent. This contradicts Condition 2 in (2.80). This contradiction proves the lemma. □

Remark 2.26 Lemma 2.25 is different from Barbalat's Lemma 2.24. While Barbalat's Lemma assumes that $f(t)$ is uniformly continuous, Lemma 2.25 assumes that $|\frac{d}{dt} f(t)|$ is bounded by $k_1 f(t)$.

Corollary 2.27 *If $f(t)$ is uniformly continuous, such that $\int_0^\infty f(\tau)\,d\tau$ exists and is finite, then $f(t) \to 0$ as $t \to \infty$.*

Corollary 2.28 *If $f(t)$, $\dot{f}(t) \in L_\infty$, and $f(t) \in L_p$, for some $p \in [1, \infty)$, then $f(t) \to 0$ as $t \to \infty$.*

Corollary 2.29 *For the differentiable function $f(t)$, if $\lim_{t\to\infty} f(t) = k < \infty$ and $\ddot{f}(t)$ exists, then $\dot{f}(t) \to 0$ as $t \to \infty$.*

Lemma 2.30 *Consider a scalar system*

$$\dot{x} = -cx + p(t) \tag{2.83}$$

where $c > 0$ and $p(t)$ is a bounded and uniformly continuous function. If, for any initial time $t_0 > 0$ and any initial condition $x(t_0)$, the solution $x(t)$ is bounded and converges to 0 as $t \to \infty$, then

$$\lim_{t\to\infty} p(t) = 0 \tag{2.84}$$

Lemma 2.31 *Consider a first-order differential equation of the form*

$$\dot{x} = -\big(a(t) + f_1\big(\xi(t)\big)\big)x + f_2\big(\xi(t)\big) \tag{2.85}$$

where f_1 and f_2 are continuous functions, and $\xi : [0, \infty) \to \mathbb{R}^m$ is a time-varying vector-valued signal that exponentially converges to zero and, for all $t \geq t_0 \geq 0$, satisfies

$$\left| f_i\big(\xi(t)\big)\right| \leq \gamma_i\big(\left\|\xi(t_0)\right\|\big)e^{-\sigma_i(t-t_0)} \tag{2.86}$$

where σ_i, $i = 1, 2$ and γ_i are class-K functions. If $a(t)$ enjoys the property that there is a constant σ_3 such that

$$\int_{t_1}^{t_2} a(\tau)\,d\tau \geq \sigma_3(t_2 - t_1), \quad \forall t_2 \geq t_1 \geq 0 \tag{2.87}$$

then there exist a class-K function γ and a constant $\sigma > 0$ such that

$$\left|x(t)\right| \leq \gamma\big(\left\|x\big(t_0, \xi(t_0)\big)\right\|\big)e^{-\sigma(t-t_0)} \tag{2.88}$$

2.9 Controllability and Observability of Nonlinear Systems

2.9.1 Controllability

This section deals with the controllability and observability properties of nonlinear systems described by linear time varying state–space representations. In particular, consider a nonlinear system defined by the state–space representation:

$$\dot{x}(t) = f(x) + \sum_{i=1}^{m} g_i(x)u_i, \quad x \in \Omega_x \subset \mathbb{R}^n \tag{2.89}$$

where $u = [u_1, u_2, \ldots, u_n]^T \in \Omega_u \subset \mathbb{R}^m$ is the input vector. The system (2.89) is defined to be controllable if there exists an admissible input vector $u(t)$ such that the state $x(t)$ can converge from an initial point $x(t_0) = x_0 \in \Omega_x$ to the final point $x(t_f) \in \Omega_x$ within a finite time interval $t_f - t_0$. The controllability means that the control system is with a set of input channels through which the input can excite the states effectively to converge to the destination x_f. Then, the controllability of (2.89) should mainly depend on the function forms of all $f(x)$ and $g_i(x)$. The controllability of the nonlinear system (2.89) is based on a useful mathematical concept called Lie algebra, which is defined as follows.

Definition 2.32 A Lie algebra over the real field $\mathbb{R}$ or the complex field $\mathbb{C}$ is a vector space $\mathbb{G}$ for which a bilinear map $(X, Y) \to [X, Y]$ is defined from $\mathbb{G} \times \mathbb{G} \to \mathbb{G}$ such that

$$[X, Y] = -[Y, X], \quad X, Y \in \mathbb{G} \tag{2.90}$$

$$\big[X, [Y, Z]\big] + \big[Y, [Z, X]\big] + \big[Z, [X, Y]\big] = 0, \quad X, Y, Z \in \mathbb{G} \tag{2.91}$$

From the above definition, a Lie algebra is a vector space where an operator $[\cdot]$ is installed, which is called a Lie bracket, can be defined arbitrarily as long as it satisfies two conditions (2.90) and (2.91) simultaneously. The condition (2.90) is often called a skew symmetric relation and obviously implies that $[X, Y] = 0$. The condition (2.91) is called the Jacobi identity, which reveals a closed loop cyclic relation among any three elements in a Lie algebra.

Define a special Lie algebra $\mathcal{E}$ that collects all n-dimensional differentiable vector fields in $\mathbb{R}^n$ along with a commutative derivative relation: For any two vector fields f and $g \in \mathbb{R}^n$, which are functions of $x \in \mathbb{R}^n$, we have

$$[f, g] = \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g \tag{2.92}$$

We can see the above equation satisfy the two conditions (2.90) and (2.91) of a Lie algebra.

It is easy to extend the above Lie bracket between two vector fields to higher order derivatives, a more compact notation may be defined based on *an adjoint operator*, that is, $[f, g] = \mathrm{ad}_f g$. This new notation treats the Lie bracket $[f, g]$ as vector field g operated on by an adjoint operator $\mathrm{ad}_f = [f, \cdot]$. Therefore, for an n-order Lie bracket ($n > 1$), one can simply write

$$\big[f, \dots [f, g] \dots\big] = \mathrm{ad}_f^n g \tag{2.93}$$

For a general control system given by (2.89), we define a control Lie algebra Δ, which is spanned by all up to order $(n-1)$ Lie brackets among f and g_1 through g_m as

$$\Delta = \mathrm{span}\big\{g_1, \dots, g_m, \mathrm{ad}_f g_1, \dots, \mathrm{ad}_f g_m, \dots, \mathrm{ad}_f^{n-1} g_1, \dots, \mathrm{ad}_f^{n-1} g_m\big\} \tag{2.94}$$

With the control Lie algebra concept, we can show that the following theorem is true and is also a general effective testing criterion for system controllability.

Theorem 2.33 *The control system* (2.89) *is controllable if and only if* $\dim(\Delta) = \dim(\Omega_x) = n$.

Note that because each element in Δ is a function of x, the dimension of Δ may be different from one point to another. Thus, if the preceding condition of dimension is valid only in a neighborhood of a point in $\Omega_x \subset \mathbb{R}^n$, we say that the system (2.89) is locally controllable. On the other hand, if the condition of dimension can cover all of region Ω_x, then it is globally controllable.

2.9.2 Observability

Consider the observability for the following nonlinear system

$$\dot{x} = f(x), \qquad y = h(x) \tag{2.95}$$

where $y \in \Omega_y \subset \mathbb{R}^m$ is the output vector. This system is said to be observable if for each pair of distinct states x_1 and x_2, the corresponding outputs y_1 and y_2 are also distinguishable. Clearly, the observability can be interpreted as a testing criterion to check whether the entire system has sufficient output channels to measure (or observe) each internal state change. Intuitively, the observability should depend on the function forms of both $f(x)$ and $h(x)$.

We introduce a Lie derivative, which is virtually *a directional derivative* for a scalar field $\lambda(x)$, with $x \in \mathbb{R}^n$ along the direction of an n-dimensional vector field $f(x)$. The mathematical expression is given as

$$L_f \lambda(x) = \frac{\partial \lambda(x)}{\partial x} f(x) \tag{2.96}$$

Since $\frac{\partial \lambda(x)}{\partial x}$ is a $1 \times n$ gradient vector of the scalar $\lambda(x)$ and the norm of a gradient vector represents the maximum rate of function value changes, the product of the gradient and the vector field $f(x)$ in (2.95) becomes the directional derivative of $\lambda(x)$ along $f(x)$. Therefore, the Lie derivative of a scalar field defined by (2.96) is also a scalar field. If each component of a vector field $h(x) \in \mathbb{R}^m$ is considered to take a Lie derivative along $f(x) \in \mathbb{R}^n$, then all components can be acted on concurrently and the result is a vector field that has the same dimension as $h(x)$; its ith element is the Lie derivative of the ith component of $h(x)$. Namely, if $h(x) = [h_1(x), \ldots, h_m(x)]^T$ and each component $h_i(x)$, $i = 1, \ldots, m$ is a scalar field, then the Lie derivative of the vector field $h(x)$ is defined as

$$L_f h(x) = \begin{bmatrix} L_f h_1(x) \\ \vdots \\ L_f h_m(x) \end{bmatrix} \tag{2.97}$$

With the Lie derivative concept, we now define an observation space Ω_0 over $\mathbb{R}^n$ as

$$\Omega_0 = \operatorname{span}\left\{h(x), L_f h(x), \ldots, L_f^{n-1} h(x)\right\} \tag{2.98}$$

In other words, this space is spanned by all up to order $(n-1)$ Lie derivatives of the output function $h(x)$. Then, we further define an observability distribution, denoted by $d\Omega_0$, which collects "the gradient" vector of every component in Ω_0. Namely,

$$d\Omega_0 = \operatorname{span}\left\{\frac{\partial \phi}{\partial x} \mid \phi \in \Omega_0\right\} \tag{2.99}$$

With these definitions, we can present the following theorem for testing the observability.

Theorem 2.34 *The system* (2.95) *is observable if and only if* $\dim(d\Omega_0) = n$.

Similarly to the controllability case, this testing criterion also has locally observable and globally observable cases, depending on whether the condition of dimension in the theorem is valid only in a neighborhood of a point or over the entire state region.

2.9.3 Brockett's Theorem on Feedback Stabilization

The following theorem, which is due to Brockett [21], gives a necessary condition for the existence of a stabilizing control law for the system

$$\dot{x} = f(x, u) \tag{2.100}$$

at an equilibrium point x_0 with x being the state and u being the control input.

Theorem 2.35 *Let the system* (2.100) *be given with* $f(x_0, 0) = 0$ *and* $f(x, u)$ *continuously differentiable in a neighborhood of* $(x_0, 0)$. *A necessary condition for the existence of a continuously differentiable control law that makes* $(x_0, 0)$ *asymptotically stable is that*

1. *the linearized system should have no uncontrollable modes associated with eigenvalues whose real part is positive*;
2. *there exists a neighborhood* Ω *of* $(x_0, 0)$ *such that for each* $\xi \in \Omega$ *there exists a control* u_ξ *defined on* $[0, \infty)$ *such that this control steers the solution of* $\dot{x} = f(x, u_\xi)$ *from* $x = \xi$ *at* $t = 0$ *to* $x = x_0$ *at* $t = \infty$;
3. *the mapping* $A \times \mathbb{R}^m \to \mathbb{R}^n$ *defined by* $(x, u) \to f(x, u)$ *should be onto an open set containing* 0.

Remark 2.36 If the system (2.100) is of the form

$$\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x) \tag{2.101}$$

then condition of Theorem 2.35 implies that the stabilization problem cannot have a solution if there is a smooth distribution D containing $f(\cdot)$ and $g_1(\cdot), \ldots, g_m(\cdot)$ with $\dim D < n$. One further special case: If the system (2.100) is of the form

$$\dot{x} = \sum_{i=1}^{m} u_i g_i(x), \quad x \in \Omega \subset \mathbb{R}^n \tag{2.102}$$

with the vectors $g_i(x)$ being linearly independent at x_0, then there exists a solution to the stabilization problem if and only if $m = n$. In this case, we must have as many control parameters as we have dimensions of x.

2.10 Lyapunov Theorems

The Lyapunov approach provides a rigorous method for addressing stability. The method is a generalization of the idea that if there is some "measure of energy" in a system, then we can study the rate of change of the energy of the system to ascertain stability. Here, we review several concepts that are used in Lyapunov stability theory.

Definition 2.37 A continuous function $V : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is a locally positive definite function if for some $\varepsilon > 0$ and some continuous, strictly increasing function $\alpha : \mathbb{R}_+ \to \mathbb{R}$,

$$V(0,t) = 0, \quad \text{and} \quad V(x,t) \geq \alpha\|x\|, \quad \forall t \geq 0 \tag{2.103}$$

Definition 2.38 A continuous function $V : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is a positive definite function if it satisfies the conditions of Definition 2.37 and, additionally, $\alpha(p) \to \infty$ as $p \to \infty$.

Definition 2.39 A continuous function $V : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is decrescent if for some $\varepsilon > 0$ and some continuous, strictly increasing function $\beta : \mathbb{R}_+ \to \mathbb{R}$,

$$V(x,t) \leq \beta\big(\|x\|\big), \quad \forall x \in \Omega, \ \forall t \geq 0 \tag{2.104}$$

Using these definitions, the following theorem allows us to determine stability for a system by studying an appropriate energy function. Roughly, this theorem states that when $V(x,t)$ is a locally positive definite function and $V(x,t) \leq 0$ then we can conclude stability of the equilibrium point. The time derivative of V is taken along the trajectories of the system:

$$\dot{V}(x,t) = \frac{dV}{dt}(x,t) = \frac{\partial V}{\partial t}(x,t) + \left[\frac{\partial V}{\partial x}(x,t)\right]^T f(t,x) \tag{2.105}$$

Theorem 2.40 (Lyapunov Theorem) *Any nonlinear dynamic system*

$$\dot{x} = f(x,t), \quad x(0) = x_0 \tag{2.106}$$

with an the equilibrium point at the origin, let Ω be a ball of size around the origin, i.e., $\Omega = \{x : \|x\| \leq \varepsilon, \ \varepsilon > 0\}$,

(1) *if for $x \in \Omega$, there exists a scalar function $V(x,t) > 0$ such that $-\dot{V}(x,t) \leq 0$, the origin of system is stable*;
(2) *if for $x \in \Omega$, there exists a scalar decrescent function $V(x,t) > 0$ such that $-\dot{V}(x,t) \leq 0$, the origin of system is uniformly stable*;
(3) *if for $x \in \Omega$, there exists a scalar function $V(x,t) > 0$ such that $-\dot{V}(x,t) \leq 0$, the origin of system is asymptotically stable*;
(4) *if for $x \in \Omega$ there exists a scalar decrescent function $V(x,t) > 0$ such that $-\dot{V}(x,t) \leq 0$, the origin of system is uniformly asymptotically stable*;
(5) *if and only if there exists an $\varepsilon > 0$ and a function $V(x,t)$ which satisfies $\gamma_1\|x\|^2 \leq V(x,t) \leq \gamma_2\|x\|^2$, $\dot{V} \leq -\gamma_3\|x\|^2$, $\|\frac{\partial V}{\partial x}(x,t)\| \leq \gamma_4\|x\|$ with some positive constants γ_1, γ_2, γ_3, γ_4 and $\|x\| \leq \varepsilon$, $x = 0$ is an exponentially stable equilibrium point of $x = f(x,t)$*;
(6) *if there exist a scalar decrescent function $V : \mathbb{R}^n \times \mathbb{R}_+ \mapsto \mathbb{R}$, and a time $t_0 \geq 0$, such that $\dot{V} \leq 0$, and for any sufficiently small positive real number r, there exists a non-origin point $x \in \mathbf{B}_r$ such that $V(x,t_0) \geq 0$, the origin of system is unstable.*

The function $V(x,t)$ is called the Lyapunov function.

The indirect method of Lyapunov uses the linearization of a system to determine the local stability of the original system.

Theorem 2.41 (Stability by Linearization) *Consider the system $\dot{x} = f(x,t)$ and define*

$$A(t) = \frac{\partial f(x,t)}{\partial x} \tag{2.107}$$

with $x = 0$ to be the Jacobian matrix of $f(x,t)$ with respect to x, evaluated at the origin. It follows that for each fixed t,

$$f_1(x,t) = f(x,t) - A(t)x \tag{2.108}$$

approaches zero as x approaches zero. Assume

$$\lim_{\|x\|\to 0} \sup_{t\geq 0} \frac{\|f_1(x,t)\|}{\|x\|} = 0 \tag{2.109}$$

Further, if 0 is a uniformly asymptotically stable equilibrium point of

$$\dot{z} = A(t)z \tag{2.110}$$

then it is a locally uniformly asymptotically stable equilibrium point of $\dot{x} = f(x,t)$.

Invariant Set Theorems Asymptotic stability of a control system is a very important property. However, the Lyapunov theorems are usually difficult to apply because frequently $\dot{V}$, the derivative of the Lyapunov function candidate, is only semi-definite. With the help of the invariant set theorems, asymptotic stability can still possibly be concluded for autonomous systems from LaSalle's invariance principle [113]. The concept of an invariant set is a generalization of the concept of equilibrium point.

Definition 2.42 (α Limit Set) The set $\Omega \in \mathbb{R}^n$ is the α limit set of a trajectory $\omega(t, x_0, t_0)$ if for every $y \in \Omega$, there exists a strictly increasing sequence of times T such that $\omega(T, x_0, t_0) \to y$ as $T \to \infty$.

Definition 2.43 A set $\Omega \in \mathbb{R}^n$ is said to be an invariant set of the dynamic system $\dot{x} = f(x)$ if for all $y \in \Omega$ and $t_0 > 0$, we have $\omega(t, y, t_0) \in \Omega, \forall t > t_0$.

Theorem 2.44 (LaSalle's Theorem) *Let Ω be a compact invariant set $\Omega = x \in \mathbb{R}^n : \dot{V}(x) = 0$ and $V : \mathbb{R}^n \to \mathbb{R}$ be a locally positive definite function such that on the compact set we have $\dot{V}(x) \leq 0$. As $t \to \infty$, the trajectory tends to the largest invariant set inside Ω; i.e., its α limit set is contained inside the largest invariant set in Ω. In particular, if Ω contains no invariant sets other than $x = 0$, then 0 is asymptotically stable.*

Corollary 2.45 *Given the autonomous nonlinear system*

$$\dot{x} = f(x), \quad x(0) = x_0 \tag{2.111}$$

and let the origin be an equilibrium point, $V(x) : \mathcal{N} \to \mathbb{R}$ be a continuously differentiable positive definite function on a neighborhood $\mathcal{N}$ of the origin, such that $\dot{V}(x) \leq 0$ in $\mathcal{N}$, then the origin is asymptotically stable if there is no solution that can stay forever in $S = x \in \mathcal{N} \mid \dot{V}(x) = 0$, other than the trial solution. The origin is globally asymptotically stable if $\mathcal{N} = \mathbb{R}^n$ and $V(x)$ is radially unbounded.

Other Stability Results

Definition 2.46 A system is said to be BIBO (bounded-input and bounded-output) stable iff for any bounded-input, the output is bounded, i.e., if for any

$$\|u\| < M < \infty \tag{2.112}$$

there exist finite $\alpha > 0$ and β such that

$$\|y\| \leq \alpha M + \beta \tag{2.113}$$

Theorem 2.47 *Let the closed-loop transfer function $H(s) \in \mathbb{R}^{n \times n}(s)$ be exponentially stable and strictly proper, and $h(t)$ be the corresponding impulse response. We have the following properties*:

(1) *For any minimal representation of $H(s)$ (i.e., any minimal state representation of the form $\dot{x} = Ax + Bu$, $y = Cx$ with $H(s) = C(sI - A)^{-1}B$), the equilibrium point $x = 0$ is globally, exponentially, uniformly (in t) stable.*
(2) *If $u \cup L_1^n$, then $y = h * u \in L_1^n \cap L_\infty^n$, $\dot{y} \in L_\infty^n$, y is absolutely continuous and $y(t) \to 0$ as $t \to \infty$.*
(3) *If $u \in L_2^n$, then $y = h * u \in L_2^n \cap L_\infty^n$, $\dot{y} \in L_2^n$, y is continuous and $y(t) \to 0$ as $t \to \infty$.*
(4) *If $u \in L_\infty^n$, then $y = h * u \in L_\infty^n \cap L_\infty^n$, $\dot{y} \in L_\infty^n$, y is continuous and $y(t) \to 0$ as $t \to \infty$.*
(5) *If $u \in L_\infty^n$, and if $u(t) \to u_\infty$ (a constant vector in $\mathbb{R}$) as $t \to \infty$, then $y \to H(0)u_\infty$, as $t \to \infty$ and the convergence is exponential.*
(6) *If $u \in L_p^n$ and $1 < p < \infty$, then, $y = h * u \in L_p^n$ and $\dot{y} \in L_p^n$.*

$h * u$ denotes the convolution product of h and u.

Lemma 2.48 *Let $e(t) = h * r$, where $h = L^{-1}(H(s))$ and $H(s)$ is an $n \times n$ strictly proper, exponentially stable transfer function. Then $r \in L_2^n \Rightarrow e \in L_2^n \cap L_\infty^n$, $\dot{e} \to L_2^n$, e is continuous and $e \to 0$ as $t \to \infty$. If, in addition, $r \to 0$ as $t \to \infty$, then $\dot{e} \to 0$ [32].*

Based on this lemma, we have the following corollaries, which are used in the proof of the stability properties of the closed-loop system [44].

Corollary 2.49 *If $H(s)$ is an exponentially stable and strictly proper diagonal matrix with*

$$H(s) \triangleq \operatorname{diag} \frac{s^m + n_i(s)}{s^{m+p} + d_i(s)} \tag{2.114}$$

where $n_i(s)$ and $d_i(s)$ are the remaining polynomial terms of s for the ith entry, then, $s^i H(s)$, $1 \le i < p$ are exponentially stable and strictly proper, too. Therefore, if $r \in L_2^n \Rightarrow e, \dot{e}, \dots, e^{(p-1)} \in L_2^n \cap L_\infty^n$, and $e, \dot{e}, \dots, e^{(p-1)}$ are continuous and $e, \dot{e}, \dots, e^{(p-1)} \to 0$ as $t \to \infty$, then $e^{(p)} \to 0$.

Corollary 2.50 *If $H(s)$ is defined as above and $p = 2$, then, $sH(s)$ is exponentially stable and strictly proper, too. Therefore, $r \in L_2^n \Rightarrow e$ and $\dot{e} \in L_2^n \cap L_\infty^n$, $\ddot{e} \in L_2^n$, e and $\dot{e} \in L_2^n \cap L_\infty^n$ are continuous and e and $\dot{e} \to 0$ as $t \to \infty$. In addition, if $r \to 0$ as $t \to \infty$, then $\ddot{e} \to 0$.*

Theorem 2.51 *Given a linear time-invariant (LTI) system*

$$\dot{x}(t) = Ax(t) \tag{2.115}$$

the system is stable iff there exists a symmetric positive definite solution P_L to the Lyapunov equation

$$A^T P_L + P_L A = -Q_L \tag{2.116}$$

where Q_L is an arbitrary symmetric positive-definite matrix.

Theorem 2.52 *For a given stable LTI system*

$$\dot{x}(t) = Ax(t) \tag{2.117}$$

the system is exponentially uniformly stable.

Definition 2.53 A function $T : U_o \to \mathbb{R}^n$ is called a diffeomorphism if it is smooth, and if its inverse, T^{-1} exists and is smooth. If the region U_o is the whole space $\mathbb{R}^n$, then $T(x)$ is called a global diffeomorphism [113].

Definition 2.54 A nonlinear system

$$\dot{x} = f(x) + \sum_{i=1}^{n} g_i(x)u_i \tag{2.118}$$

$$= f(x) + G(x)u \tag{2.119}$$

is said to be feedback linearizable in a neighborhood U_o of the origin if there exist a diffeomorphism $T : U_o \to \mathbb{R}^n$ and a nonlinear feedback control law

$$u = M^{-1}(x)\big[v - F(x)\big] \tag{2.120}$$

such that the transformed state

$$z = T(x) \tag{2.121}$$

and the new input v satisfies the linear time invariant system

$$\dot{z} = Az + Bv \tag{2.122}$$

where (A, B) is a controllable linear system [113].

Consider the linear time-invariant system given by

$$\dot{x}(t) = Ax(t), \quad t \geq 0 \tag{2.123}$$

In this special case, Lyapunov theory is very complete, and we have the following theorem.

Theorem 2.55 *For linear time-invariant system* (2.123), *the following statements are equivalent*:

(i) *the system is asymptotically stable*;
(ii) *the system is exponential stable*; *matrix A is Hurwitz*;
(iv) *the Lyapunov equation*

$$A^T P + PA = -Q \tag{2.124}$$

has a unique solution $P > 0$ for any $Q > 0$; *and*
(v) *Eq*. (2.124) *has a unique solution $P > 0$ for some $Q > 0$*.

The last statement asserts the existence of a quadratic Lyapunov function

$$V(x) = x^T Px$$

where P is symmetric and positive definite.

For the discrete-time linear time-invariant system

$$x_{k+1} = Ax_k, \quad k \in \mathbf{N}_+ \tag{2.125}$$

a similar result may be stated as follows.

2.11 Notes and References

This chapter briefly provides the fundamental concepts and tools that will be used for control design and stability analysis of wheeled inverted pendulum in the coming chapters. The interested reader is referred to [67, 72, 74, 79, 81, 95] for a detailed and comprehensive coverage of the topics discussed in this chapter.

http://www.springer.com/978-1-4471-2962-2

Advanced Control of Wheeled Inverted Pendulum Systems
Li, Z.; Yang, C.; Fan, L.
2013, XIV, 218 p., Hardcover
ISBN: 978-1-4471-2962-2