2

Background Material

This chapter consists primarily of some background material, with the selection of topics being dictated by our later needs. Some facts and structural concepts of the linear case have a marked parallel in MJLS, so they are included here in order to facilitate the comparison. In Section 2.1 we introduce the notation, norms, and spaces that are appropriate for our approach. Next, in Section 2.2, we present some important auxiliary results that will be used throughout the book. In Section 2.3 we discuss some issues on the probability space for the underlined model. In Sections 2.4 and 2.5, we recall some basic facts regarding linear systems and linear matrix inequalities.

2.1 Some Basics

We shall use throughout the book some standard definitions and results from operator theory in Banach spaces which can be found, for instance, in [181] or [216]. For $\mathbb{X}$ and $\mathbb{Y}$ complex Banach spaces we set $\mathbb{B}(\mathbb{X}, \mathbb{Y})$ the Banach space of all bounded linear operators of $\mathbb{X}$ into $\mathbb{Y}$, with the uniform induced norm represented by $\|.\|$. For simplicity we set $\mathbb{B}(\mathbb{X}) \triangleq \mathbb{B}(\mathbb{X},\mathbb{X})$. The spectral radius of an operator $\mathcal{T} \in \mathbb{B}(\mathbb{X})$ is denoted by $r_\sigma(\mathcal{T})$. If $\mathbb{X}$ is a Hilbert space then the inner product is denoted by $\langle .;. \rangle$, and for $\mathcal{T} \in \mathbb{B}(\mathbb{X})$, $\mathcal{T}^*$ denotes the adjoint operator of $\mathcal{T}$. As usual, $\mathcal{T} \geq 0$ ($\mathcal{T} > 0$ respectively) will denote that the operator $\mathcal{T} \in \mathbb{B}(\mathbb{X})$ is positive-semi-definite (positive-definite). In particular, we denote respectively by $\mathbb{R}^n$ and $\mathbb{C}^n$ the n dimensional real and complex Euclidean spaces and $\mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$ ($\mathbb{B}(\mathbb{R}^n, \mathbb{R}^m)$ respectively) the normed bounded linear space of all $m \times n$ complex (real) matrices, with $\mathbb{B}(\mathbb{C}^n) \triangleq \mathbb{B}(\mathbb{C}^n, \mathbb{C}^n)$ ($\mathbb{B}(\mathbb{R}^n) \triangleq \mathbb{B}(\mathbb{R}^n, \mathbb{R}^n)$). Unless otherwise stated, $\|.\|$ will denote the standard norm in $\mathbb{C}^n$, and for $M \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$, $\|M\|$ denotes the induced uniform norm in $\mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$. The superscript * indicates the conjugate transpose of a matrix, while $'$ indicates the transpose. Clearly for real matrices * and $'$ will have the same meaning. The identity operator is denoted by $\mathcal{I}$, and the $n \times n$ identity matrix by I_n (or simply I). Finally, we denote by $\lambda_i(P)$, $i = 1, \ldots, n$ the eigenvalues of a matrix $P \in \mathbb{B}(\mathbb{C}^n)$.

Remark 2.1. We recall that the trace operator $\operatorname{tr}(.) : \mathbb{B}(\mathbb{C}^n) \to \mathbb{C}$ is a linear functional with the following properties:

1. $\operatorname{tr}(KL) = \operatorname{tr}(LK)$. (2.1a)
2. For any $M, P \in \mathbb{B}(\mathbb{C}^n)$ with $M \geq 0, P > 0$,

$$\left(\min_{i=1,\ldots,n} \lambda_i(P)\right) \operatorname{tr}(M) \leq \operatorname{tr}(MP) \leq \left(\max_{i=1,\ldots,n} \lambda_i(P)\right) \operatorname{tr}(M). \quad (2.1b)$$

In this book we shall be dealing with finite dimensional spaces, in which case all norms are equivalent. It is worth recalling that two norms $\|.\|_1, \|.\|_2$ in a Banach space $\mathbb{X}$ are equivalent if for some $c_1 > 0$ and $c_2 > 0$, and all $x \in \mathbb{X}$,

$$\|x\|_1 \leq c_2 \|x\|_2 \,,\; \|x\|_2 \leq c_1 \|x\|_1 \,.$$

As we are going to see in the next chapters, to analyze the stochastic model as in (1.3), we will use the indicator function on the jump parameter to *markovianize* the state. This, in turn, will decompose the matrices associated to the second moment and control problems into N matrices. Therefore it comes up naturally that a convenient space to be used is the one we define as $\mathbb{H}^{n,m}$, which is the linear space made up of all N-sequences of complex matrices $V = (V_1, \ldots, V_N)$ with $V_i \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$, $i \in \mathbb{N}$. For simplicity, we set $\mathbb{H}^n \triangleq \mathbb{H}^{n,n}$. For $V = (V_1, \ldots, V_N) \in \mathbb{H}^{n,m}$, we define the following equivalent norms in the finite dimensional space $\mathbb{H}^{n,m}$:

$$\begin{aligned} \|V\|_1 &\triangleq \sum_{i=1}^{N} \|V_i\| \\ \|V\|_2 &\triangleq \left(\sum_{i=1}^{N} \operatorname{tr}(V_i^* V_i)\right)^{1/2} \\ \|V\|_{\max} &\triangleq \max\{\|V_i\|\,;\, i \in \mathbb{N}\}. \end{aligned} \quad (2.2)$$

We shall omit the subscripts $1, 2, \max$ whenever the definition of a specific norm does not affect the result being considered. It is easy to verify that $\mathbb{H}^{n,m}$ equipped with any of the above norms is a Banach space and, in fact, $(\|.\|_2, \mathbb{H}^{n,m})$ is a Hilbert space, with the inner product given, for $V = (V_1, \ldots, V_N)$ and $S = (S_1, \ldots, S_N)$ in $\mathbb{H}^{n,m}$, by

$$\langle V; S\rangle \triangleq \sum_{i=1}^{N} \operatorname{tr}(V_i^* S_i). \quad (2.3)$$

It is also convenient to define the following equivalent induced norms $\|.\|_1$ and $\|.\|_2$ in the finite dimensional space $\mathbb{B}(\mathbb{H}^n)$. For $\mathcal{T} \in \mathbb{B}(\mathbb{H}^n)$,

$$\|\mathcal{T}\|_1 \triangleq \sup_{V\in\mathbb{H}^n} \frac{\|\mathcal{T}(V)\|_1}{\|V\|_1}, \quad \|\mathcal{T}\|_2 \triangleq \sup_{V\in\mathbb{H}^n} \frac{\|\mathcal{T}(V)\|_2}{\|V\|_2}.$$

Again, we shall omit the subscripts $1, 2$ whenever the definition of the specific norm does not matter to the problem under consideration. For $V = (V_1, \ldots, V_N) \in \mathbb{H}^{n,m}$ we write $V^* = (V_1^*, \ldots, V_N^*) \in \mathbb{H}^{m,n}$ and say that $V \in \mathbb{H}^n$ is hermitian if $V = V^*$. We set

$$\mathbb{H}^{n*} \triangleq \{V = (V_1, \ldots, V_N) \in \mathbb{H}^n; V_i = V_i^*,\, i \in \mathbb{N}\}$$

and

$$\mathbb{H}^{n+} \triangleq \{V = (V_1, \ldots, V_N) \in \mathbb{H}^{n*}; V_i \geq 0,\, i \in \mathbb{N}\}$$

and write, for $V = (V_1, \ldots, V_N) \in \mathbb{H}^n$ and $S = (S_1, \ldots, S_N) \in \mathbb{H}^n$, that $V \geq S$ if $V - S = (V_1 - S_1, \ldots, V_N - S_N) \in \mathbb{H}^{n+}$, and that $V > S$ if $V_i - S_i > 0$ for $i \in \mathbb{N}$. We say that an operator $\mathcal{T} \in \mathbb{B}(\mathbb{H}^n)$ is hermitian if $\mathcal{T}(V) \in \mathbb{H}^{n*}$ whenever $V \in \mathbb{H}^{n*}$, and that it is positive if $\mathcal{T}(V) \in \mathbb{H}^{n+}$ whenever $V \in \mathbb{H}^{n+}$.

We define the operators φ and $\hat{\varphi}$ in the following way: for $V = (V_1, \ldots, V_N) \in \mathbb{H}^{n,m}$, with $V_i = \begin{bmatrix} v_{i1} & \cdots & v_{in} \end{bmatrix} \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^m), v_{ij} \in \mathbb{C}^m$

$$\varphi(V_i) \triangleq \begin{bmatrix} v_{i1} \\ \vdots \\ v_{in} \end{bmatrix} \in \mathbb{C}^{mn} \quad \text{and} \quad \hat{\varphi}(V) \triangleq \begin{bmatrix} \varphi(V_1) \\ \vdots \\ \varphi(V_N) \end{bmatrix} \in \mathbb{C}^{Nmn}.$$

With the Kronecker product $L \otimes K \in \mathbb{B}(\mathbb{C}^{n^2})$ defined in the usual way for any $L, K \in \mathbb{B}(\mathbb{C}^n)$, the following properties hold (see [43]):

$$(L \otimes K)^* = L^* \otimes K^* \tag{2.4a}$$

$$\varphi(LKH) = (H' \otimes L)\varphi(K), H \in \mathbb{B}(\mathbb{C}^n). \tag{2.4b}$$

Remark 2.2. It is easy to verify, through the mapping $\hat{\varphi}$, that the spaces $\mathbb{H}^{n,m}$ and $\mathbb{C}^{Nmn}$ are uniformly homeomorphic (see [181], p. 117) and that any operator $\mathcal{Z}$ in $\mathbb{B}(\mathbb{H}^{n,m})$ can be represented in $\mathbb{B}(\mathbb{C}^{Nmn})$ through the mapping $\hat{\varphi}$. We shall denote this operator by $\hat{\varphi}[\mathcal{Z}]$. Clearly we must have

$$r_\sigma(\mathcal{Z}) = r_\sigma(\hat{\varphi}[\mathcal{Z}]).$$

Remark 2.3. It is well known that if $W \in \mathbb{B}(\mathbb{C}^n)^+$ then there exists a unique $W^{1/2} \in \mathbb{B}(\mathbb{C}^n)^+$ such that $W = (W^{1/2})^2$. The absolute value of $W \in \mathbb{B}(\mathbb{C}^n)$, denoted by $|W|$, is defined as $|W| = (W^*W)^{1/2}$. As shown in [216], p. 170, there exists an orthogonal matrix $U \in \mathbb{B}(\mathbb{C}^n)$ (that is, $U^{-1} = U^*$) such that

$$W = U\,|W| \ \ (\text{ or } \ |W| = U^{-1}W = U^*W), \tag{2.5}$$

and $\|W\| = \|\,|W|\,\|$.

Remark 2.4. For any $W \in \mathbb{B}(\mathbb{C}^n)$ there exist W^j, $j = 1, 2, 3, 4$, such that $W^j \geq 0$ and $\|W^j\| \leq \|W\|$ for $j = 1, 2, 3, 4$, and $W = (W^1 - W^2) + \sqrt{-1}(W^3 - W^4)$. Indeed, we can write

$$W = V^1 + \sqrt{-1}V^2$$

where

$$V^1 = \frac{1}{2}(W^* + W)$$
$$V^2 = \frac{\sqrt{-1}}{2}(W^* - W).$$

Since V^1 and V^2 are self-adjoint (that is, $V^i = V^{i*}, i = 1, 2$), and every self-adjoint element in $\mathbb{B}(\mathbb{C}^n)$ can be decomposed into positive and negative parts (see [181], p. 464), we have that there exist $W^i \in \mathbb{B}(\mathbb{C}^n)^+$, $i = 1, 2, 3, 4$, such that

$$V^1 = W^1 - W^2$$
$$V^2 = W^3 - W^4.$$

Therefore for any $S = (S_1, \ldots, S_N) \in \mathbb{H}^n$, we can find $S^j \in \mathbb{H}^{n+}$, $j = 1, 2, 3, 4$ such that $\|S^j\|_1 \leq \|S\|_1$ and

$$S = (S^1 - S^2) + \sqrt{-1}(S^3 - S^4).$$

2.2 Auxiliary Results

The next result follows from the decomposition of square matrices into positive semi-definite matrices as seen in Remark 2.4 in conjunction with Lemma 1 and Remark 4 in [156].

Proposition 2.5. *Let $\mathcal{Z} \in \mathbb{B}(\mathbb{H}^n)$. The following assertions are equivalent:*

1. $\sum_{k=0}^{\infty} \|\mathcal{Z}^k(V)\|_1 < \infty$ *for all* $V \in \mathbb{H}^{n+}$.
2. $r_\sigma(\mathcal{Z}) < 1$.
3. $\|\mathcal{Z}^k\| \leq \beta\zeta^k, k = 0, 1, \ldots$ *for some* $0 < \zeta < 1$ *and* $\beta \geq 1$.
4. $\|\mathcal{Z}^k(V)\|_1 \to 0$ *as* $k \to \infty$ *for all* $V \in \mathbb{H}^{n+}$.

Proof. From Remark 2.4 for any $S = (S_1, \ldots, S_N) \in \mathbb{H}^n$, we can find $S^j \in \mathbb{H}^{n+}$, $j = 1, 2, 3, 4$ such that $\|S^j\|_1 \leq \|S\|_1$ and

$$S = (S^1 - S^2) + \sqrt{-1}(S^3 - S^4).$$

Since $\mathcal{Z}$ is a linear operator we get

$$\begin{aligned}\left\|\mathcal{Z}^k(S)\right\|_1 &= \left\|\mathcal{Z}^k(S^1) - \mathcal{Z}^k(S^2) + \sqrt{-1}\left(\mathcal{Z}^k(S^3) - \mathcal{Z}^k(S^4)\right)\right\|_1 \\ &\leq \sum_{i=1}^{4} \left\|\mathcal{Z}^k(S^i)\right\|_1 .\end{aligned} \tag{2.6}$$

The result now follows easily from Lemma 1 and Remark 4 in [156], after noticing that $\mathbb{H}^n$ is a finite dimensional complex Banach space. □

The next result is an immediate adaptation of Lemma 1 in [158].

Proposition 2.6. *Let $\mathcal{Z} \in \mathbb{B}(\mathbb{H}^n)$. If $r_\sigma(\mathcal{Z}) < 1$ then there exists a unique $V \in \mathbb{H}^n$ such that*

$$V = \mathcal{Z}(V) + S$$

for any $S \in \mathbb{H}^n$. Moreover,

$$V = \hat{\varphi}^{-1}\left(\left(I_{Nn^2} - \hat{\varphi}[\mathcal{Z}]\right)^{-1} \hat{\varphi}(S)\right) = \sum_{k=0}^{\infty} \mathcal{Z}^k(S).$$

Furthermore, if $\mathcal{Z}$ is a hermitian operator then

$$S = S^* \Leftrightarrow V = V^*,$$

and if $\mathcal{Z}$ is a positive operator then

$$\begin{aligned} S \geq 0 \ &\Rightarrow V \geq 0 \\ S > 0 \ &\Rightarrow V > 0. \end{aligned}$$

The following corollary is an immediate consequence of the previous result.

Corollary 2.7. *Suppose that $\mathcal{Z} \in \mathbb{B}(\mathbb{H}^n)$ is a positive operator with $r_\sigma(\mathcal{Z}) < 1$. If*

$$\begin{aligned} V &= \mathcal{Z}(V) + S \\ \tilde{V} &= \mathcal{Z}(\tilde{V}) + \tilde{S}, \end{aligned}$$

with $\tilde{S} \geq S$ ($\tilde{S} > S$) then $\tilde{V} \geq V$ ($\tilde{V} > V$).

Proof. Straightforward from Proposition 2.6. □

The following definition and result will be useful in Chapter 3.

Definition 2.8. *We shall say that a Cauchy sequence $\{z(k); k = 0, 1, \ldots\}$ in a complete normed space $\mathbb{Z}$ (in particular, $\mathbb{C}^n$ or $\mathbb{B}(\mathbb{C}^n)$) is Cauchy summable if (cf. [157])*

$$\sum_{k=0}^{\infty} \sup_{\tau \geq 0} \|z(k+\tau) - z(k)\| < \infty.$$

The next proposition was established in Lemma (L1) of [157].

Proposition 2.9. *Let $\{z(k); k = 0, 1, \ldots\}$ be a Cauchy summable sequence in $\mathbb{Z}$ and consider the sequence $\{y(k); k = 0, 1, \ldots\}$ in $\mathbb{Z}$ given by*

$$y(k + 1) = \mathcal{L}y(k) + z(k)$$

where $\mathcal{L} \in \mathbb{B}(\mathbb{Z})$. If $r_\sigma(\mathcal{L}) < 1$, then $\{y(k); k = 0, 1, \ldots\}$ is a Cauchy summable sequence and for any initial condition $y(0) \in \mathbb{Z}$,

$$\lim_{k\to\infty} y(k) = (\mathcal{I} - \mathcal{L})^{-1} \lim_{k\to\infty} z(k).$$

2.3 Probabilistic Space

In this section we present in detail the probabilistic framework we shall consider throughout this book. We shall be dealing with stochastic models as in (1.3) with, at each time k, the jump variable $\theta(k)$ taking values in the set $\mathbb{N} = \{1, \ldots, N\}$, and the remaining input variables taking values in $\tilde{\Omega}_k$. Thus, for the jump variable, we set $\mathfrak{N}$ the σ-field of all subsets of $\mathbb{N}$, and for the remaining input variables, we set $\tilde{\mathfrak{F}}_k$ as the Borel σ-field of $\tilde{\Omega}_k$. To consider all time values, we define

$$\Omega \triangleq \prod_{k\in\mathbb{T}} (\tilde{\Omega}_k \times \mathbb{N}_k)$$

where $\mathbb{N}_k$ are copies of $\mathbb{N}$, $\times$ and $\prod$ denote the product space, and $\mathbb{T}$ represents the discrete-time set, being $\{\ldots, -1, 0, 1, \ldots\}$ when the process starts from $-\infty$, or $\{0, 1, \ldots\}$, when the process starts from 0. Set also $\mathbb{T}_k = \{i \in \mathbb{T}; i \le k\}$ for each $k \in \mathbb{T}$, and

$$\mathfrak{F} \triangleq \sigma \left\{ \prod_{k\in\mathbb{T}} S_k \times \psi_k ;\ S_k \in \tilde{\mathfrak{F}}_k \text{ and } \psi_k \in \mathfrak{N} \text{ for each } k \in \mathbb{T} \right\}$$

and for each $k \in \mathbb{T}$,

$$\mathfrak{F}_k \triangleq \sigma \left\{ \prod_{l\in\mathbb{T}_k} S_l \times \psi_l \times \prod_{\tau=k+1}^{\infty} \tilde{\Omega}_\tau \times \mathbb{N}_\tau ; S_l \in \tilde{\mathfrak{F}}_l \text{ and } \psi_l \in \mathfrak{N} \text{ for } l \in \mathbb{T}_k \right\}$$

so that $\mathfrak{F}_k \subset \mathfrak{F}$. We define then the stochastic basis $(\Omega, \mathfrak{F}, \{\mathfrak{F}_k\}, \mathcal{P})$, where $\mathcal{P}$ is a probability measure such that

$$\mathcal{P}(\theta(k + 1) = j \mid \mathfrak{F}_k) = \mathcal{P}(\theta(k + 1) = j \mid \theta(k)) = p_{\theta(k)j}$$

with $p_{ij} \ge 0$ for $i, j \in \mathbb{N}$, $\sum_{j=1}^{N} p_{ij} = 1$ for each $i \in \mathbb{N}$, and for each $k \in \mathbb{T}$, $\theta(k)$ is a random variable from Ω to $\mathbb{N}$ defined as $\theta(k)(\omega) = \beta(k)$ with $\omega = \{(\xi(k), \beta(k)); k \in \mathbb{T}\}$, $\xi(k) \in \tilde{\Omega}_k, \beta(k) \in \mathbb{N}$. Clearly $\{\theta(k); k \in \mathbb{T}\}$ is a Markov chain taking values in $\mathbb{N}$ and with transition probability matrix $\mathrm{P} = [p_{ij}]$. The initial distribution for $\theta(0)$ is denoted by $\upsilon = \{\upsilon_1, \ldots, \upsilon_N\}$.

We set $C^m = L_2(\Omega, \mathfrak{F}, \mathcal{P}, \mathbb{C}^m)$ the Hilbert space of all second order $\mathbb{C}^m$-valued $\mathfrak{F}$-measurable random variables with inner product given by $\langle x; y \rangle = E(x^* y)$ for all $x, y \in C^m$, where $E(.)$ stands for the expectation of the underlying scalar valued random variables, and norm denoted by $\|.\|_2$. Set $\ell_2(C^m) = \underset{k\in\mathbb{T}}{\oplus} C^m$, the direct sum of countably infinite copies of C^m, which is a Hilbert space made up of $r = \{r(k); k \in \mathbb{T}\}$, with $r(k) \in C^m$ for each $k \in \mathbb{T}$, and such that

$$\|r\|_2^2 \triangleq \sum_{k\in\mathbb{T}} E(\|r(k)\|^2) < \infty.$$

For $r = \{r(k); k \in \mathbb{T}\} \in \ell_2(C^m)$ and $v = \{v(k); k \in \mathbb{T}\} \in \ell_2(C^m)$, the inner product $\langle r; s \rangle$ in $\ell_2(C^m)$ is given by

$$\langle r; s \rangle \triangleq \sum_{k\in\mathbb{T}} E(r^*(k)v(k)) \leq \|r\|_2 \|v\|_2 .$$

We define $\mathcal{C}^m \subset \ell_2(C^m)$ in the following way: $r = \{r(k); k \in \mathbb{T}\} \in \mathcal{C}^m$ if $r \in \ell_2(C^m)$ and $r(k) \in L_2(\Omega, \mathfrak{F}_k, \mathcal{P}, \mathbb{C}^m)$ for each $k \in \mathbb{T}$. We have that $\mathcal{C}^m$ is a closed linear subspace of $\ell_2(C^m)$ and therefore a Hilbert space. We also define $\mathcal{C}_k^m$ as formed by the elements $r_k = \{r(k); k \in \mathbb{T}_k\}$ such that $r(l) \in L_2(\Omega, \mathfrak{F}_l, \mathcal{P}, \mathbb{C}^m)$ for each $l \in \mathbb{T}_k$. Finally we define Θ_0 as the set of all $\mathfrak{F}_0$-measurable variables taking values in $\mathbb{N}$.

2.4 Linear System Theory

Although MJLS seem, prima facie, a natural extension of the linear class, their subtleties are such that the standard linear theory cannot be directly applied, although it will be most illuminating in the development of the results described in this book. In view of this, it is worth having a brief look at some basic results and properties of the linear time-invariant systems (in short LTI), whose Markov jump counterparts will be considered later.

2.4.1 Stability and the Lyapunov Equation

Consider the following difference equations

$$x(k+1) = f(x(k)) \tag{2.7}$$

and

$$x(k+1) = Ax(k) \tag{2.8}$$

with $k \in \{0, 1, 2, \ldots\}$, $x(k) \in \mathbb{C}^n$, $f : \mathbb{C}^n \to \mathbb{C}^n$ and $A \in \mathbb{B}(\mathbb{C}^n)$. A sequence $x(0), x(1), \ldots$ generated according to (2.7) or (2.8) is called a trajectory of the system. The second equation is a particular case of the first one and is of greater interest to us (thus we shall not be concerned on regularity hypotheses over f in (2.7)). It defines what we call a *discrete-time homogeneous linear time-invariant system*. For more information on dynamic systems or proofs of the results presented in this section, the reader may refer to one of the many works on the theme, like [48], [165] and [213].

First we recall that a point $x_e \in \mathbb{C}^n$ is called an *equilibrium point* of System (2.7), if $f(x_e) = x_e$. In particular, $x_e = 0$ is an equilibrium point of System (2.8). The following definitions apply to Systems (2.7) and (2.8).

Definition 2.10 (Lyapunov Stability). *An equilibrium point x_e is said to be stable in the sense of Lyapunov if for each $\epsilon > 0$ there exists $\delta_\epsilon > 0$ such that $\|x(k) - x_e\| \leq \epsilon$ for all $k \geq 0$ whenever $\|x(0) - x_e\| \leq \delta_\epsilon$.*

Definition 2.11 (Asymptotic Stability). *An equilibrium point is said to be asymptotically stable if it is stable in the sense of Lyapunov and there exists $\delta > 0$ such that whenever $\|x(0) - x_e\| \leq \delta$ we have that $x(k) \to x_e$ as k increases. It is globally asymptotically stable if it is asymptotically stable and $x(k) \to x_e$ as k increases for any $x(0)$ in the state space.*

The definition above simply states that the equilibrium point is stable if, given any spherical region surrounding the equilibrium point, we can find another spherical region surrounding the equilibrium point such that trajectories starting inside this second region do not leave the first one. Besides, if the trajectories also converge to this equilibrium point, then it is asymptotically stable.

Definition 2.12 (Lyapunov Function). *Let x_e be an equilibrium point for System (2.7). A positive function $\phi : \Gamma \to \mathbb{R}$, where Γ is such that $x_e \in \Gamma \subseteq \mathbb{C}^n$, is said to be a Lyapunov function for System (2.7) and equilibrium point x_e if*

1. *$\phi(.)$ is continuous,*
2. *$\phi(x_e) < \phi(x)$ for every $x \in \Gamma$ such that $x \neq x_e$,*
3. *$\Delta\phi(x) = \phi(f(x)) - \phi(x) \leq 0$ for all $x \in \Gamma$.*

With this we can proceed to the Lyapunov Theorem. A proof of this result can be found in [165].

Theorem 2.13 (Lyapunov Theorem). *If there exists a Lyapunov function $\phi(x)$ for System (2.7) and x_e, then the equilibrium point is stable in the sense of Lyapunov. Moreover, if $\Delta\phi(x) < 0$ for all $x \neq x_e$, then it is asymptotically stable. Furthermore if ϕ is defined on the entire state space and $\phi(x)$ goes to infinity as any component of x gets arbitrarily large in magnitude then the equilibrium point x_e is globally asymptotically stable.*

The Lyapunov theorem applies to System (2.7) and, of course, to System (2.8) as well. Let us consider a possible Lyapunov function for System (2.8) as follows:

$$\phi(x(k)) = x^*(k)Vx(k) \tag{2.9}$$

with $V > 0$. Then

$$\begin{aligned}\Delta\phi(x(k)) &= \phi(x(k+1)) - \phi(x(k)) \\ &= x^*(k+1)Vx(k+1) - x^*(k)Vx(k) \\ &= x^*(k)A^*VAx(k) - x^*(k) * Vx(k) \\ &= x^*(k)\left(A^*VA - V\right)x(k).\end{aligned}$$

With this we can present the following theorem that establishes the connection between System (2.8), stability, and the so called Lyapunov equation. All assertions are classical applications of the Lyapunov theorem with the Lyapunov function (2.9). The proof can be found, for instance, in [48].

Theorem 2.14. *The following assertions are equivalent.*

1. $x = 0$ *is the only globally asymptotically stable equilibrium point for System (2.8).*
2. $r_\sigma(A) < 1$.
3. *For any* $S > 0$*, there exists a unique* $V > 0$ *such that*

$$V - A^*VA = S. \tag{2.10}$$

4. *For some* $V > 0$*, we have*

$$V - A^*VA > 0. \tag{2.11}$$

The above theorem will be extended to the Markov case in Chapter 3 (Theorem 3.9).

Since (2.8) has only one equilibrium point whenever it is stable, we commonly say in this case that System (2.8) is stable.

2.4.2 Controllability and Observability

Let us now consider a non-homogeneous form for System (2.8)

$$x(k+1) = Ax(k) + Bu(k) \tag{2.12}$$

where $B \in \mathbb{B}(\mathbb{C}^m, \mathbb{C}^n)$ and $u(k) \in \mathbb{C}^m$ is a vector of inputs to the system.

The idea behind the concept of controllability is rather simple. It deals with answering the following question: for a certain pair (A, B), is it possible to apply a sequence of $u(k)$ in order to drive the system from any $x(0)$ to a specified final state x_f in a finite time?

The following definition establishes more precisely the concept of controllability. Although not treated here, a concept akin to controllability is the *reachability* of a system. In more general situations these concepts may differ, but in the present case they are equivalent, and therefore we will only use the term controllability.

Definition 2.15 (Controllability). *The pair (A, B) is said to be* controllable*, if for any $x(0)$ and any given final state x_f, there exists a finite positive integer T and a sequence of inputs $u(0), u(1), \ldots, u(T-1)$ that, applied to System (2.12), yields $x(T) = x_f$.*

One can establish if a given system is controllable using the following theorem, which also lists some classical results (see [48], p. 288).

Theorem 2.16. *The following assertions are equivalent.*

1. *The pair (A, B) is controllable.*
2. *The following $n \times nm$ matrix (called a* controllability matrix*) has rank n:*
$$\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}.$$
3. *The* controllability Grammian *$S_c \in \mathbb{B}(\mathbb{C}^n)$ given by*
$$S_c(k) = \sum_{i=0}^{k} A^i BB^* (A^*)^i$$
is nonsingular for some $k < \infty$.
4. *For A and B real, given any monic real polynomial ψ of degree n, there exists $F \in \mathbb{B}(\mathbb{R}^n, \mathbb{R}^m)$ such that $\det(sI - (A + BF)) = \psi(s)$.*

Moreover, if $r_\sigma(A) < 1$ then the pair (A, B) is controllable if and only if the unique solution S_c of $S = ASA^* + BB^*$ is positive-definite.

The concept of controllability Grammian for MJLS will be presented in Chapter 4, Section 4.4.2.

Item 4 of the theorem above is particularly interesting, since it involves the idea of *state feedback*. Suppose that for some $F \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$, we apply $u(k) = Fx(k)$ in System (2.12), yielding
$$x(k+1) = (A + BF)x(k),$$
which is a form similar to (2.8). According to the theorem above, an adequate choice of F (for A, B and F real) would allow us to perform pole placement for the closed loop system $(A+BF)$. For instance we could use state feedback to stabilize an unstable system.

The case in which the state feedback can only change the unstable eigenvalues of the system is of great interest and leads us to the introduction of the concept of *stabilizability*.

Definition 2.17 (Stabilizability). *The pair (A, B) is said to be* stabilizable *if there exists $F \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$ such that $r_\sigma(A + BF) < 1$.*

This concept will play a crucial role for the MJLS, as will be seen in the next chapters (see Section 3.5). Consider now a system of the form

$$x(k+1) = Ax(k) \tag{2.13a}$$

$$y(k) = Lx(k) \tag{2.13b}$$

where $L \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^p)$ and $y(k) \in \mathbb{C}^p$ is the vector of outputs of the system. The concepts of controllability and stabilizability just presented, which relate structurally $x(k)$ and the input $u(k)$, have their dual counterparts from the point of view of the output $y(k)$. The following theorem and definitions present them.

Definition 2.18 (Observability). *The pair (L, A) is said to be* observable*, if there exists a finite positive integer T such that knowledge of the outputs $y(0), y(1), \ldots, y(T-1)$ is sufficient to determine the initial state $x(0)$.*

The concept of observability deals with the following question: is it possible to infer the internal behavior of a system by observing its outputs? This is a fundamental property when it comes to control and filtering issues.

The following theorem is dual to Theorem 2.16, and the proof can be found in [48], p. 282.

Theorem 2.19. *The following assertions are equivalent.*

1. *The pair (L, A) is observable.*
2. *The following $pn \times n$ matrix (called an* observability matrix*) has rank n:*

$$\begin{bmatrix} L \\ LA \\ \vdots \\ LA^{n-1} \end{bmatrix}.$$

3. *The* observability Grammian *$S_o \in \mathbb{B}(\mathbb{C}^n)$ given by*

$$S_o(k) = \sum_{i=0}^{k} (A^*)^i L^* L A^i$$

   is nonsingular for some $k < \infty$.
4. *For A and L real, given any monic real polynomial ψ of degree n, there exists $K \in \mathbb{B}(\mathbb{R}^p, \mathbb{R}^n)$ such that* $\det(sI - (A + KL)) = \psi(s)$.

Moreover, if $r_\sigma(A) < 1$ then the pair (L, A) is observable if and only if the unique solution S_o of $S = A^* S A + L^* L$ is positive-definite.

We also define the concept of *detectability*, which is dual to the definition of stabilizability.

Definition 2.20 (Detectability). *The pair (L, A) is said to be* detectable *if there exists $K \in \mathbb{B}(\mathbb{C}^p, \mathbb{C}^n)$ such that $r_\sigma(A + KL) < 1$.*

These are key concepts in linear system theory which will be extended to the Markov jump case in due course throughout this book.

2.4.3 The Algebraic Riccati Equation and the Linear-Quadratic Regulator

Consider again System (2.12)

$$x(k+1) = Ax(k) + Bu(k).$$

An extensively studied and classical control problem is that of finding a sequence $u(0), u(1), \ldots, u(T-1)$ that minimizes the cost $\mathfrak{J}_T(x_0, u)$ given by

$$\mathfrak{J}_T(x_0, u) = \sum_{k=0}^{T-1} \left[\|Cx(k)\|^2 + \|Du(k)\|^2 \right] + E(x(T)^* \mathcal{V} x(T)), \tag{2.14}$$

where $\mathcal{V} \geq 0$ and $D^*D > 0$. The idea of minimizing $\mathfrak{J}_T(x_0, u)$ is to drive the state of the system to the origin without much strain from the control variable which is, in general, a desirable behavior for control systems. This problem is referred to as the *linear-quadratic regulator* (*linear* system + *quadratic* cost) problem. It can be shown (see for instance [48] or [183]) that the solution to this problem is

$$u(k) = F(k)x(k) \tag{2.15}$$

with $F(k)$ given by

$$\begin{aligned} F(k) &= -(B^* X_T(k+1) B + D^* D)^{-1} B^* X_T(k+1) A & \text{(2.16a)} \\ X_T(k) &= C^* C + A^* X_T(k+1) A - A^* X_T(k+1) B \\ &\quad \times (B^* X_T(k+1) B + D^* D)^{-1} B^* X_T(k+1) A & \text{(2.16b)} \\ X_T(T) &= \mathcal{V}. \end{aligned}$$

Equation (2.16b) is called the *difference Riccati equation*. Another related problem is the infinite horizon linear quadratic regulator problem, in which it is desired to minimize the cost

$$\mathfrak{J}(x_0, u) = \sum_{k=0}^{\infty} \left[\|Cx(k)\|^2 + \|Du(k)\|^2 \right]. \tag{2.17}$$

Under some conditions, the solution to this problem is

$$u(k) = F(X)x(k), \tag{2.18}$$

where the constant gain $F(X)$ is given by

$$F(X) = -(B^*XB + D^*D)^{-1}B^*XA \tag{2.19}$$

and X is a positive semi-definite solution of

$$W = C^*C + A^*WA - A^*WB(B^*WB + D^*D)^{-1}B^*WA. \tag{2.20}$$

Equation (2.20) is usually referred to as the *algebraic Riccati equation* or in short, ARE . If $r_\sigma(A+BF(X)) < 1$, then X is said to be a stabilizing solution of (2.20). Questions that naturally arise are: under which conditions there is convergence of $X_T(0)$ given by (2.16b) as T goes to infinity to a positive semi-definite solution X of (2.20)? When is there a stabilizing solution for (2.20)? Is it unique? The following theorem, whose proof can be found, for instance, in [48], p. 348, answers these questions.

Theorem 2.21. *Suppose that the pair (A, B) is stabilizable. Then for any $\mathcal{V} \geq 0$, $X_T(0)$ converges to a positive semi-definite solution X of (2.20) as T goes to infinity. Moreover if the pair (C, A) is detectable, then there exists a unique positive semi-definite solution X to (2.20), and this solution is the unique stabilizing solution for (2.20).*

Riccati equations like (2.16b) and (2.20) and their variations are employed in a variety of control (as in (2.14) and (2.17)) and filtering problems. As we are going to see in Chapters 4, 5, 6 and 7, they will also play a crucial role for MJLS. For more on Riccati equations and associated problems, see [26], [44], and [195].

2.5 Linear Matrix Inequalities

Some miscellaneous definitions and results involving matrices and matrix equations are presented in this section. These results will be used throughout the book, especially those related with the concept of linear matrix inequalities (or in short LMIs), which will play a very important role in the next chapters.

Definition 2.22 (Generalized Inverse). *The generalized inverse (or Moore–Penrose inverse) of a matrix $A \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$ is the unique matrix $A^\dagger \in \mathbb{B}(\mathbb{C}^m, \mathbb{C}^n)$ such that*

1. $AA^\dagger A = A$,
2. $A^\dagger AA^\dagger = A^\dagger$,
3. $(AA^\dagger)^* = AA^\dagger$,
4. $(A^\dagger A)^* = A^\dagger A$.

For more on this subject, see [49]. The Schur complements presented below are used to convert quadratic equations into larger dimension linear ones and vice versa.

Lemma 2.23 (Schur complements). *(From [195]). Consider an hermitian matrix Q such that*

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^* & Q_{22} \end{bmatrix}.$$

1. $Q > 0$ if and only if

$$\begin{cases} Q_{22} > 0 \\ Q_{11} - Q_{12}Q_{22}^{-1}Q_{12}^* > 0 \end{cases}$$

or

$$\begin{cases} Q_{11} > 0 \\ Q_{22} - Q_{12}^*Q_{11}^{-1}Q_{12} > 0. \end{cases}$$

2. $Q \geq 0$ if and only if

$$\begin{cases} Q_{22} \geq 0 \\ Q_{12} = Q_{12}Q_{22}^{\dagger}Q_{22} \\ Q_{11} - Q_{12}Q_{22}^{\dagger}Q_{12}^* \geq 0 \end{cases}$$

or

$$\begin{cases} Q_{11} \geq 0 \\ Q_{12} = Q_{11}Q_{11}^{\dagger}Q_{12} \\ Q_{22} - Q_{12}^*Q_{11}^{\dagger}Q_{12} \geq 0\ . \end{cases}$$

Next we present the definition of LMI.

Definition 2.24. *A linear matrix inequality (LMI) is any constraint that can be written or converted to*

$$F(x) = F_0 + x_1F_1 + x_2F_2 + \ldots + x_mF_m < 0, \tag{2.21}$$

where x_i are the variables and the hermitian matrices $F_i \in \mathbb{B}(\mathbb{R}^n)$ for $i = 1, \ldots, m$ are known.

LMI (2.21) is referred to as a strict LMI. Also of interest are the nonstrict LMIs, where $F(x) \leq 0$. From the practical point of view, LMIs are usually presented as

$$f(X_1, \ldots, X_N) < g(X_1, \ldots, X_N), \tag{2.22}$$

where f and g are affine functions of the unknown matrices $X_1, \ldots, X_N$. For example, from the Lyapunov equation, the stability of System (2.8) is equivalent to the existence of a $V > 0$ satisfying the LMI (2.11). Quadratic forms can usually be converted to affine ones using the Schur complements. Therefore we will make no distinctions between (2.21) and (2.22), quadratic and affine forms, or between a set of LMIs or a single one, and will refer to all of them as simply LMIs. For more on LMIs the reader is referred to [7], [42], or any of the many works on the subject.

http://www.springer.com/978-1-85233-761-2

Discrete-Time Markov Jump Linear Systems
Costa, O.L.V.; Fragoso, M.D.; Marques, R.P.
2005, X, 286 p. 15 illus., Hardcover
ISBN: 978-1-85233-761-2