David L. Elliott

APPLIED MATHEMATICAL SCIENCES 169

Bilinear Control Systems

Matrices in Action

Springer

Applied Mathematical Sciences

Volume 169

For further volumes:
http://www.springer.com/series/34

This page intentionally left blank

David L. Elliott

Bilinear Control Systems

Matrices in Action

Springer

Prof. David Elliott
University of Maryland
Inst. Systems Research
College Park MD 20742
USA

Editors:

ISSN 0066-5452
ISBN 978-1-4020-9612-9 e-ISBN 978-1-4020-9613-6
DOI 10.1007/978-1-4020-9613-6
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009920095

Mathematics Subject Classification (2000): 93B05, 1502, 57R27, 22E99, 37C10

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The mathematical theory of control became a field of study half a century ago in attempts to clarify and organize some challenging practical problems and the methods used to solve them. It is known for the breadth of the mathematics it uses and its cross-disciplinary vigor. Its literature, which can be found in Section 93 of *Mathematical Reviews*, was at one time dominated by the theory of linear control systems, which mathematically are described by linear differential equations forced by additive control inputs. That theory led to well-regarded numerical and symbolic computational packages for control analysis and design.

Nonlinear control problems are also important; in these either the underlying dynamical system is nonlinear or the controls are applied in a non-additive way. The last four decades have seen the development of theoretical work on nonlinear control problems based on differential manifold theory, nonlinear analysis, and several other mathematical disciplines. Many of the problems that had been solved in linear control theory, plus others that are new and distinctly nonlinear, have been addressed; some resulting general definitions and theorems are adapted in this book to the bilinear case.

A nonlinear control system is called bilinear if it is described by linear differential equations in which the control inputs appear as coefficients. Such multiplicative controls (valves, interest rates, switches, catalysts, etc.) are common in engineering design and also are used as models of natural phenomena with variable growth rates. Their study began in the late 1960s and has continued from its need in applications, as a source of understandable examples in nonlinear control, and for its mathematical beauty. Recent work connects bilinear systems to such diverse disciplines as switching systems, spin control in quantum physics, and Lie semigroup theory; the field needs an expository introduction and a guide, even if incomplete, to its literature.

The control of continuous-time bilinear systems is based on properties of matrix Lie groups and Lie semigroups. For that reason, much of the first half of the book is based on matrix analysis including the Campbell–Baker–Hausdorff Theorem. (The usual approach would be to specialize geometric methods of nonlinear control based on Frobenius's Theorem in manifold theory.) Other topics such as discrete-time systems, observability and realization, applications, linearization of nonlinear systems with finite-dimensional Lie algebras, and input–output analysis have chapters of their own.

The intended readers for this book includes graduate mathematicians and engineers preparing for research or application work. If short and helpful, proofs are given; otherwise they are sketched or cited from the mathematical literature. The discussions are amplified by examples, exercises, and also a few Mathematica scripts to show how easily software can be written for bilinear control problems.

Before you begin to read, please turn to the very end of the book to see that the Index lists in bold type the pages where symbols and concepts are first defined. Then please glance at the Appendices A–D; they are referred to often for standard facts about matrices, Lie algebras, and Lie groups. Throughout the book, sections whose titles have an asterisk (*) invoke less familiar mathematics — come back to them later. The mark □ indicates the end of a proof, while △ indicates the end of a definition, remark, exercise, or example.

My thanks go to A. V. Balakrishnan, my thesis director, who introduced me to bilinear systems; my Washington University collaborators in bilinear control theory — William M. Boothby, James G-S. Cheng, Tsuyoshi Goka, Ellen S. Livingston, Jackson L. Sedwick, Tzyh-Jong Tarn, Edward N. Wilson, and the late John Zaborsky; the National Science Foundation, for support of our research; Michiel Hazewinkel, who asked for this book; Linus Kramer; Ron Mohler; Luiz A. B. San Martin; Michael Margaliot, for many helpful comments; and to the Institute for Systems Research of the University of Maryland for its hospitality since 1992. This book could not have existed without the help of Pauline W. Tang, my wife.

College Park, Maryland, *David L. Elliott*

Contents

This page intentionally left blank

Chapter 1
Introduction

Most engineers and many mathematicians are familiar with linear time-invariant control systems; a simple example can be written as a set of first-order differential equations

$$\dot{x} = Ax + u(t)b,$$

where $x \in \mathbb{R}^n$, A is a square matrix, u is a locally integrable function and $b \in \mathbb{R}^n$ is a constant vector. The idea is that we have a dynamical system that left to itself would evolve on $\mathbb{R}^n$ as $dx/dt = Ax$, and a control term $u(t)b$ can be added to influence the evolution. Linear control system theory is well established, based mostly on linear algebra and the geometry of linear spaces.

For many years, these systems and their kin dominated control and communication analysis. Problems of science and technology, as well as advances in nonlinear analysis and differential geometry, led to the development of nonlinear control system theory. In an important class of nonlinear control systems, the control u is used as a multiplicative coefficient,

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

where f and g are differentiable vector functions. The theory of such systems can be found in several textbooks such as Sontag [249] or Jurdjevic [147]. They include a class of control systems in which $f(x) := Ax$ and $g(x) := Bx$, linear functions, so

$$\dot{x} = Ax + u(t)Bx,$$

which is called a bilinear control system. (The word *bilinear* means that the velocity contains a ux term but is otherwise linear in x and u.) This specialization leads to a simpler and satisfying, if still incomplete, theory with many applications in science and engineering.

D.L. Elliott, *Bilinear Control Systems*, Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_1,

Contents of this Chapter

We begin with an important special case that does not involve control: linear dynamical systems. Sections 1.1 and 1.2 discuss them, especially their stability properties, as well as some topics in matrix analysis. The concept of *control system* will be introduced in Section 1.3 through linear control systems — with which you may be familiar. Bilinear control systems themselves are first encountered in Sections 1.4–1.6. Section 1.7 returns to the subject of stability with introductions to Lyapunov's direct method and Lyapunov exponents. Section 1.8 has comments on time-discretization issues. The exercises in Section 1.9 and scattered through the chapter are intended to illustrate and extend the main results. Note that newly defined terms are displayed in sans-serif typeface.

1.1 Matrices in Action

This section defines linear dynamical systems, then discusses their properties and some relevant matrix functions. The matrix algebra notation, terminology, and basic facts needed here are summarized in Section A.1 of Appendix A. For instance, a statement in which the symbol $\mathbb{F}$ appears is supposed to be true whether $\mathbb{F}$ is the real field $\mathbb{R}$ or the complex field $\mathbb{C}$.

1.1.1 Linear Dynamical Systems

Definition 1.1. A dynamical system on $\mathbb{F}^n$ is a triple $(\mathbb{F}^n, \mathcal{T}, \Theta)$. Here $\mathbb{F}^n$ is the state space of the dynamical system; elements (column vectors) $x \in \mathbb{F}^n$ are called states. $\mathcal{T}$ is the system's time-set (either $\mathbb{R}_+$ or $\mathbb{Z}_+$). Its transition mapping Θ_t (also called its evolution function) gives the state at time t as a mapping continuous in the initial state $x(0) := \xi \in \mathbb{F}^n$,

$$x(t) = \Theta_t(\xi),\ \Theta_0(\xi) = \xi; \text{ and}$$
$$\Theta_s(\Theta_t(\xi)) = \Theta_{s+t}(\xi) \tag{1.1}$$

for all $s, t \in \mathcal{T}$ and $\xi \in \mathbb{F}^n$. The dynamical system is called linear if for all $x, z \in \mathbb{F}^n$, $t \in \mathcal{T}$ and scalars $\alpha, \beta \in \mathbb{F}$

$$\Theta_t(\alpha x + \beta z) = \alpha\Theta_t(x) + \beta\Theta_t(z). \tag{1.2}$$

△

In this definition, if $\mathcal{T} := \mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ we call the triple a discrete-time dynamical system; its mappings Θ_t can be obtained from the mapping

$\Theta_1 : \mathbb{F}^n \to \mathbb{F}^n$ by using (1.1). Discrete-time linear dynamical systems on $\mathbb{F}^n$ have a transition mapping Θ_t that satisfies (1.2); then there exists a square matrix A such that $\Theta_1(x) = Ax$;

$$x(t+1) = Ax(t), \quad \text{or succinctly } \overset{+}{x} = Ax; \quad \Theta_t(\xi) = A^t\xi. \tag{1.3}$$

Here $\overset{+}{x}$ is called the successor of x.

If $\mathcal{T} = \mathbb{R}_+$ and Θ_t is differentiable in x and t, the triple is called a continuous-time dynamical system. Its transition mapping can be assumed to be (as in Section B.3.2) a *semi-flow* $\Theta : \mathbb{R}_+ \times \mathbb{F}^n \to \mathbb{F}^n$. That is, $x(t) = \Theta_t(\xi)$ is the unique solution, defined and continuous on $\mathbb{R}_+$, of some first-order differential equation $\dot{x} = f(x)$ with $x(0) = \xi$. This differential equation will be called the dynamics[1] and in context specifies the dynamical system.

If instead of the semi-flow assumption we postulate that the transition mapping satisfies (1.2) (linearity) then the triple is called a continuous-time linear dynamical system and

$$\lim_{t\downarrow 0} \frac{1}{t}(\Theta_t(x) - x) = Ax$$

for some matrix $A \in \mathbb{F}^{n\times n}$. From that and (1.1), it follows that the mapping $x(t) = \Theta_t(x)$ is a semi-flow generated by the dynamics

$$\dot{x} = Ax. \tag{1.4}$$

Given an initial state $x(0) := \xi$ set $x(t) := X(t)\xi$; then (1.4) reduces to a single initial-value problem for matrices: find an $n \times n$ matrix function $X : \mathbb{R}_+ \to \mathbb{F}^{n\times n}$ such that

$$\dot{X} = AX, \ X(0) = I. \tag{1.5}$$

Proposition 1.1. *The initial-value problem* (1.5) *has a unique solution* $X(t)$ *on* $\mathbb{R}_+$ *for any* $A \in \mathbb{F}^{n\times n}$.

Proof. The formal power series in t

$$X(t) := I + At + \frac{A^2t^2}{2} + \cdots + A^k\frac{t^k}{k!} + \cdots \tag{1.6}$$

and its term-by-term derivative $\dot{X}(t)$ formally satisfy (1.5). Using the inequalities (see (A.3) in Appendix A) satisfied by the operator norm $\|\cdot\|$, we see that on any finite time interval $[0, T]$

$$\|X(t) - (I + At + \cdots + \frac{A^kt^k}{k!})\| \leq \sum_{i=k+1}^{\infty} \frac{|T|^i\|A\|^i}{i!} \xrightarrow[k\to\infty]{} 0,$$

[1] Section B.3.2 includes a discussion of nonlinear dynamical systems with *finite escape time;* for them (1.1) makes sense only for sufficiently small $s + t$.

so (1.6) converges uniformly on $[0, T]$. The series for the derivative $\dot{X}(t)$ converges in the same way, so $X(t)$ is a solution of (1.5); the series for higher derivatives also converge uniformly on $[0, T]$, so $X(t)$ has derivatives of all orders.

The uniqueness of $X(t)$ can be seen by comparison with any other solution $Y(t)$. Let

$$Z(t) := X(t) - Y(t); \ \ \text{then } Z(t) = \int_0^t AZ(s)\,ds.$$

$$\text{Let } \ \zeta(t) := \|Z(t)\|, \ \alpha := \|A\|; \text{ then } \ \zeta(t) \le \int_0^t \alpha\zeta(s)\,ds$$

which implies $\zeta(t) = 0$. □

As a corollary, given $\xi \in \mathbb{F}^n$ the unique solution of the initial-value problem for (1.4) given ξ is the transition mapping $\Theta_t x = X(t)x$.

Definition 1.2. Define the matrix exponential function by $e^{tA} = X(t)$, where $X(t)$ is the solution of (1.5). △

With this definition, the transition mapping for (1.4) is given by $\Theta_t(x) = \exp(tA)x$, and (1.1) implies that the matrix exponential function has the property

$$e^{tA}e^{sA} = e^{(t+s)A}, \tag{1.7}$$

which can also be derived from the series (1.6); that is left as an exercise in series manipulation. Many ways of calculating the matrix exponential function $\exp(A)$ are described in Moler and Van Loan [211].

Exercise 1.1. For any $A \in \mathbb{F}^{n\times n}$ and integer $k \ge n$, A^k is a polynomial in A of degree $n - 1$ or less (see Appendix A). △

1.1.2 Matrix Functions

Functions of matrices like $\exp A$ can be defined in several ways. Solving a matrix differential equation like (1.5) is one way; the following power series method is another. (See Remark A.1 for the mapping from formal power series to polynomials in A.) Denote the set of eigenvalues of $A \in \mathbb{F}^{n\times n}$ by $\mathrm{spec}(A) = \{\alpha_1, \dots, \alpha_n\}$, as in Section A.1.4. Suppose a function $\psi(z)$ has a Taylor series convergent in a disc $U := \{z \mid |z| < R\}$,

$$\psi(z) = \sum_{i=0}^{\infty} c_i z^i, \text{ then define } \psi(A) := \sum_{i=0}^{\infty} c_i A^i \tag{1.8}$$

for any A such that $\mathrm{spec}(A) \subset U$. Such a function ψ will be called good at A because the series for $\psi(A)$ converges absolutely. If ψ, like polynomials and the exponential function, is an *entire function* (analytic everywhere in $\mathbb{C}$) then ψ is good at any A.[2]

Choose $T \in \mathbb{C}^{n\times n}$ such that $\hat{A} = T^{-1}AT$ is upper triangular;[3] then so is $\psi(\hat{A})$, whose eigenvalues $\psi(\alpha_i)$ are on its diagonal. Since $\mathrm{spec}(\hat{A}) = \mathrm{spec}(A)$, if ψ is good at A then

$$\mathrm{spec}(\psi(A)) = \{\psi(\alpha_1), \dots, \psi(\alpha_n)\} \text{ and } \det(\psi(A)) = \prod_{i=1}^{n} \psi(\alpha_i).$$

If $\psi = \exp$, then

$$\det(\mathrm{e}^A) = \mathrm{e}^{\mathrm{tr}(A)}, \tag{1.9}$$

which is called Abel's relation. If ψ is good at A, $\psi(A)$ is an element of the associative algebra $\{I, A\}_{\mathfrak{A}}$ generated by I and A (see Section A.2). Like any other element of the algebra, by Theorem A.1 (Cayley–Hamilton) $\psi(A)$ can be written as a polynomial in A. To find this polynomial, make use of the recursion given by the minimum polynomial $m_A(s) := s^\kappa + c_{\kappa-1}s^{\kappa-1} + \cdots + c_0$, which is the polynomial of least degree κ such that $m_A(A) = 0$.

Proposition 1.2. *If $m_A(s) = (s - \alpha_1) \cdots (s - \alpha_\kappa)$ then*

$$\mathrm{e}^{tA} = f_0(t)I + f_1(t)A + \cdots + f_{\kappa-1}(t)A^{\kappa-1} \tag{1.10}$$

where the f_i are linearly independent and of exponential polynomial form

$$f_i(t) = \sum_{k=1}^{\kappa} \mathrm{e}^{\alpha_k t} p_{ik}(t), \tag{1.11}$$

that is, the functions p_{ik} are polynomials in t. If A is real then the f_i are real.

Proof.[4] Since A satisfies $m_A(A) = 0$, e^{tA} has the form (1.10). Now operate on (1.10) with $m_A\left(\frac{d}{dt}\right)$; the result is

$$\left(c_0 I + c_1 A + \cdots + A^\kappa\right)\mathrm{e}^{tA} = m_A\left(\frac{d}{dt}\right)\left(f_0(t)I + \cdots + f_{\kappa-1}(t)A^\kappa\right).$$

Since $m_A(A) = 0$, the left hand side is zero. The matrices $I, \dots, A^{\kappa-1}$ are linearly independent, so each of the f_i satisfies the same linear time-invariant differential equation

---

[2] Matrix functions "good at A" including the matrix logarithm of Theorem A.3 are discussed carefully in Horn and Johnson [132, Ch. 6].

[3] See Section A.2.3. Even if A is real, if its eigenvalues are complex its triangular form $\hat{A}$ is complex.

[4] Compare Horn and Johnson [132, Th. 6.2.9].

$$m_A\left(\frac{d}{dt}\right)f_i = 0, \text{ and at } t = 0, \text{ for } 1 \le i, j \le \kappa \quad \frac{d^j}{dt^j}f_i(0) = \delta_{i,j}, \tag{1.12}$$

where $\delta_{i,j}$ is the Kronecker delta. Linear constant-coefficient differential equations like (1.12) have solutions $f_i(t)$ that are of the form (1.11) (exponential polynomials) and are therefore entire functions of t. Since the solution of (1.5) is unique, (1.10) has been proved. That the f_i are linearly independent functions of t follows from (1.12). □

Definition 1.3. Define the matrix logarithm of $Z \in \mathbb{F}^{n\times n}$ by the series

$$\log(Z) := (Z - I) - \frac{1}{2}(Z - I)^2 + \frac{1}{3}(Z - I)^3 - \cdots \tag{1.13}$$

which converges for $\|Z - I\| < 1$. △

1.1.3 The Λ Functions

For any real t the integral $\int_0^t \exp(s\tau)\,d\tau$ is entire. There is no established name for this function, but let us call it

$$\Lambda(s;t) := s^{-1}(\exp(ts) - 1).$$

$\Lambda(A;t)$ is good at all A, whether or not A has an inverse, and satisfies the initial-value problem

$$\frac{d\Lambda}{dt} = A\Lambda + I,\ \Lambda(A;0) = 0.$$

The continuous-time affine dynamical system $\dot{x} = Ax + b$, $x(0) = \xi$ has solution trajectories $x(t) = e^{tA}\xi + \Lambda(A;t)b$. Note:

$$\text{If } e^{TA} = I \text{ then } \Lambda(A;T) = 0. \tag{1.14}$$

The discrete-time analog of Λ is the polynomial

$$\Lambda_d(s;k) := s^{k-1} + \cdots + 1 = (s - 1)^{-1}(s^k - 1),$$

good at any matrix A. For the discrete-time affine dynamical system $\overset{+}{x} = Ax + b$ with $x(0) = \xi$ and any A, the solution sequence is $x(t) = A^t\xi + \Lambda_d(A;t)b,\ t \in N$.

$$\text{If } A^T = I \text{ then } \Lambda_d(A;T) = 0. \tag{1.15}$$

1.2 Stability: Linear Dynamics

Qualitatively, the most important property of $\dot{x} = Ax$ is the stability or instability of the equilibrium solution $x(t) = 0$, which is easy to decide in this case. (Stability questions for nonlinear dynamical systems will be examined in Section 1.7.)

Matrix $A \in \mathbb{F}^{n\times n}$ is said to be a Hurwitz matrix if there exists $\epsilon > 0$ such that $\Re(\alpha_i) < -\epsilon$, $i \in 1, \cdots, n$. For that A, there exists a positive constant k such that $\|x(t)\| < k\|\xi\| \exp(-\epsilon t)$ for all ξ, and the origin is called exponentially stable.

If the origin is not exponentially stable but $\|\exp(tA)\|$ is bounded as $t \to \infty$, the equilibrium solution $x(t) = 0$ is called stable. If also $\|\exp(-tA)\|$ is bounded as $t \to \infty$, the equilibrium is neutrally stable which implies that all of the eigenvalues are imaginary ($\alpha_k = i\omega_k$). The solutions are oscillatory; they are periodic if and only if the ω_k are all integer multiples of some ω_0. The remaining possibility for linear dynamics (1.4) is that $\Re(\alpha) > 0$ for some $\alpha \in \operatorname{spec}(A)$; then almost all solutions are unbounded as $t \to \infty$ and the equilibrium at 0 is said to be unstable.

Remark 1.1. Even if all n eigenvalues are imaginary, neutral stability is not guaranteed: if the Jordan canonical form of A has off-diagonal entries in the Jordan block for eigenvalue $\alpha = i\omega$, $x(t)$ has terms of the form $t^m \cos(\omega t)$ (resonance terms) that are unbounded on $\mathbb{R}_+$. The proof is left to Exercise 1.4, Section 1.9. △

Rather than calculate eigenvalues to check stability, we will use Lyapunov's direct (second) method, which is described in the monograph of Hahn [113] and recent textbooks on nonlinear systems such as Khalil [156]. Some properties of symmetric matrices must be described before stating Proposition 1.4.

If Q is symmetric or Hermitian (Section A.1.2), its eigenvalues are real and the number of eigenvalues of Q that are positive, zero, and negative are respectively denoted by $\{\mathfrak{p}_Q, \mathfrak{z}_Q, \mathfrak{n}_Q\}$; this list of nonnegative integers can be called the sign pattern of Q. (There are $(n+1)(n+2)/2$ possible sign patterns.) Such a matrix Q is called positive definite (one writes $Q \gg 0$) if $\mathfrak{p}_Q = n$; in that case $x^\top Qx > 0$ for all nonzero x. Q is called negative definite, written $Q \ll 0$, if $\mathfrak{n}_Q = n$.

Certain of the kth-order *minors* of a matrix Q (Section A.1.2) are its leading principal minors, the determinants

$$D_1 = q_{1,1}, \ldots, D_k = \begin{vmatrix} q_{1,1} & \cdots & q_{1,k} \\ \cdots & \cdots & \cdots \\ q_{1,k} & \cdots & q_{k,k} \end{vmatrix}, \ldots, D_n = \det(Q). \tag{1.16}$$

Proposition 1.3 (J. J. Sylvester). *Suppose $Q^* = Q$, then $Q \gg 0$ if and only if $D_i > 0$, $1 \le i \le n$.*

Proofs are given in Gantmacher [101, Ch. X, Th. 3] and Horn and Johnson [131].

If $\operatorname{spec}(A) = \{\alpha_1, \dots, \alpha_n\}$ the linear operator

$$\mathrm{Ly}_A : \mathrm{Symm}(n) \to \mathrm{Symm}(n), \quad \mathrm{Ly}_A(Q) := A^\top Q + QA,$$

called the Lyapunov operator, has the n^2 eigenvalues $\alpha_i + \alpha_j$, $1 \le i, j \le n$, so $\mathrm{Ly}_A(Q)$ is invertible if and only if all the sums $\alpha_i + \alpha_j$ are non-zero.

Proposition 1.4 (A. M. Lyapunov). *The real matrix A is a Hurwitz matrix if and only if there exist matrices $P, Q \in \mathrm{Symm}(n)$ such that $Q \gg 0$, $P \gg 0$ and*

$$\mathrm{Ly}_A(Q) = -P. \tag{1.17}$$

For proofs of this proposition, see Gantmacher [101, Ch. XV], Hahn [113, Ch. 4], or Horn and Johnson [132, 2.2].

The test in Proposition 1.4 fails if any eigenvalue is purely imaginary; but even then, if P belongs to the range of Ly_A then (1.17) has solutions Q with $\mathfrak{z}_Q > 0$.

The Lyapunov equation (1.17) is a special case of $AX + XB = C$, the Sylvester equation; the operator $X \to AX + XB$, where $A, B, X \in \mathbb{F}^{n\times n}$, is called the Sylvester operator[5] and is discussed in Section A.3.3.

Remark 1.2. If $A \in \mathbb{R}^{n\times n}$ its complex eigenvalues occur in conjugate pairs. For any similar matrix $C := S^{-1}AS \in \mathbb{C}^n$, such as a triangular (Section A.2.3) or Jordan canonical form, both C and A are Hurwitz matrices if for any Hermitian $P \gg 0$ the Lyapunov equation $C^*Q + QC = -P$ has a Hermitian solution $Q \gg 0$. △

1.3 Linear Control Systems

Linear control systems[6] of a special type provide a motivating example, permit making some preliminary definitions that can be easily extended, and are a context for facts that we will use later. The difference between a control system and a dynamical system is freedom of choice. Given an initial state ξ, instead of a single solution one has a family of them. For example, most vehicles can be regarded as control systems steered by humans or computers to desired orientations and positions in $\mathbb{R}^2$ or $\mathbb{R}^3$.

---

[5] The Sylvester equation is studied in Bellman [25], Gantmacher [101, Ch. VIII], and Horn and Johnson [132, 2.2]. It can be solved numerically in MatLab, Mathematica, etc.; the best-known algorithm is due to Bartels and Stewart [24].

[6] A general concept of *control system* is treated by Sontag [249, Ch. 2]. There are many books on linear control systems suitable for engineering students, including Kailath [152] and Corless and Frazho [65].

A (continuous-time) linear control system can be defined as a quadruple $(\mathbb{F}^n, \mathbb{R}_+, \mathcal{U}, \Theta)$ where $\mathbb{F}^n$ is the state space,[7] the time-set is $\mathbb{R}_+$, and $\mathcal{U}$ is a class of input functions $u : \mathbb{R}_+ \to \mathbb{R}^m$ (also called controls). The transition mapping Θ is parametrized by $u \in \mathcal{U}$: it is generated by the controlled dynamics

$$\dot{x}(t) = Ax(t) + Bu(t),\ u \in \mathcal{U}, \tag{1.18}$$

which in context is enough to specify the control system. There may optionally be an output $y(t) = Cx(t) \in \mathbb{F}^p$ as part of the control system's description; see Chapter 5. We will assume time-invariance, meaning that the coefficient matrices $A \in \mathbb{F}^{n\times n}$, $B \in \mathbb{F}^n \times \mathbb{R}^m$, $C \in \mathbb{R}^p \times \mathbb{F}^n$ are constant.

The largest class $\mathcal{U}$ we will need is $\mathcal{LI}$, the class of $\mathbb{R}^m$-valued locally integrable functions — those u for which the Lebesgue integral $\int_s^t \|u(\tau)\|\, d\tau$ exists whenever $0 \le s \le t < \infty$. To get an explicit transition mapping, use the fact that on any time interval $[0, T]$ a control $u \in \mathcal{LI}$ and an initial state $x(0) = \xi$ determine the unique solution of (1.18) given by

$$x(t) = e^{tA}\xi + \int_0^t e^{(t-\tau)A} Bu(\tau)\, d\tau,\ 0 \le t \le T. \tag{1.19}$$

The specification of the control system commonly includes a control constraint $u(\cdot) \in \Omega$ where Ω is a closed set. Examples are $\Omega := \mathbb{R}^m$ or $\Omega := \{u \,|\, |u_i(t)| \le 1, 1 \le i \le m\}$.[8] The input function space $\mathcal{U} \subset \mathcal{LI}$ must be invariant under time-shifts.

Definition 1.4 (Shifts). The function v obtained by shifting a function u to the right along the time axis by a finite time $\sigma \ge 0$ is written $v = \mathrm{S}^\sigma u$ where $\mathrm{S}^{(\cdot)}$ is the time-shift operator defined as follows. If the domain of u is $[0, T]$, then

$$\mathrm{S}^\sigma u(t) = \begin{cases} 0, & t \in [0, \sigma) \\ u(t-\sigma), & t \in [\sigma, T+\sigma]. \end{cases} \qquad \triangle$$

Often $\mathfrak{U}$ will be $\mathcal{PK}$, the space of piecewise constant functions, which is a shift-invariant subspace of $\mathcal{LI}$. The defining property of $\mathcal{PK}$ is that given some interval $[0, T]$, u is defined and takes on constant values in $\mathbb{R}^m$ on all the open intervals of a partition of $[0, T]$. (At the endpoints of the intervals u need not be defined.) Another shift-invariant subspace of $\mathcal{LI}$ (again defined with respect to partitions) is $\mathcal{PC}[0, T]$, the $\mathbb{R}^m$-valued piecewise continuous functions on $[0, T]$; it is assumed that the limits at the endpoints of the pieces are finite.

With an output $y = Cx$ and $\xi = 0$ in (1.19), the relation between input and output becomes a linear mapping

[7] State spaces linear over $\mathbb{C}$ are needed for computations with triangularizations throughout, and in Section 6.5 for quantum mechanical systems.

[8] Less usually, Ω may be a finite set, for instance $\Omega = \{-1, 1\}$, as in Section 6.3.

$$y = \mathbf{L}u \text{ where } \mathbf{L}u(t) := C\int_0^t e^{(t-\tau)A}Bu(\tau)\,ds,\ 0 \le t \le \infty. \tag{1.20}$$

Since the coefficients A, B, C are constant, it is easily checked that $S^\sigma \mathbf{L}(u) = \mathbf{L}S^\sigma(u)$. Such an operator $\mathbf{L}$ is time-invariant; $\mathbf{L}$ is also causal: $\mathbf{L}u(T)$ depends only on the input's past history $\mathfrak{U}_T := \{u(s), s < T\}$.

Much of the theory and practice of control systems involves the way inputs are constructed. In the study of a linear control system (1.18), a control given as a function of time is called an open-loop control, such as a sinusoidal input $u(t) := \sin(\omega t)$ used as a test signal. This is in contrast to a linear state feedback $u(t) = Kx(t)$ in which the matrix K is chosen to give the resulting system, $\dot{x} = (A + BK)x$, some desired dynamical property. In the design of a linear system to respond to an input v, commonly the system's output $y(t) = C(t)$ is used in an output feedback term,[9] so $u(t) = v(t) - \alpha Cy(t)$.

1.4 What Is a Bilinear Control System?

An exposition of control systems often would proceed at this point in either of two ways. One is to introduce nonlinear control systems with dynamics

$$\dot{x} = f(x) + \sum_1^m u_i(t)g_i(x), \tag{1.21}$$

where $f, g_1, \ldots, g_m$ are continuously differentiable mappings $\mathbb{R}^n \to \mathbb{R}^n$ (better, call them vector fields as in Section B.3.1).[10] The other usual way would be to discuss time-variant linear control systems.[11] In this book we are interested in a third way: what happens if in the linear time-variant equation $\dot{x} = F(t)x$ we let $F(t)$ be a finite linear combination of constant matrices with arbitrary time-varying coefficients?

That way leads to the following definition. Given constant matrices $A, B_1, \ldots, B_m$ in $\mathbb{R}^{n\times n}$, and controls $u \in \mathcal{U}$ with $u(t) \in \Omega \subset \mathbb{R}^m$,

$$\dot{x} = Ax + \sum_1^m u_i(t)B_ix \tag{1.22}$$

will be called a bilinear control system on $\mathbb{R}^n$.

Again the abbreviation $u := \mathrm{col}(u_1, \ldots, u_m)$ will be convenient, and as will the list $\mathbf{B}^m := \{B_1, \ldots B_m\}$ of control matrices. The term Ax is called the

[9] In many control systems, u is obtained from electrical circuits that implement causal functionals $\{y(\tau), \tau \le t\} \mapsto u(t)$.

[10] Read Sontag [249] to explore that way, although we will touch on it in Chapters 3 and 7.

[11] See Antsaklis and Michel [7].

drift term. Given any of our choices of $\mathcal{U}$, the differential equation (1.22) becomes linear time-variant and has a unique solution that satisfies (1.22) almost everywhere. A generalization of (1.22), useful beginning with Chapter 3, is to give control u as a function $\phi : \mathbb{R}^n \to \mathbb{R}^m$; $u = \phi(x)$ is then called a feedback control.[12] In that case, the existence and uniqueness of solutions to (1.22) require some technical conditions on ϕ that will be addressed when the subject arises.

Time-variant bilinear control systems $\dot{x} = A(t)x + u(t)B(t)x$ have been discussed in Isidori and Ruberti [142] and in works on system identification, but are not covered here.

Bilinear systems without drift terms and with symmetric constraints[13]

$$\dot{x} = \Big(\sum_{1}^{m} u_i(t)B_i\Big)x, \quad \Omega = -\Omega, \tag{1.23}$$

are called symmetric bilinear control systems. Among bilinear control systems, they have the most complete theory, introduced in Section 1.5.3 and Chapter 2. A drift term Ax is just a control term u_0B_0x with $B_0 = A$ and the constant control $u_0 = 1$. The control theory of systems with drift terms is more difficult and incomplete; it is the topic of Chapter 3.

Bilinear systems were introduced in the U.S. by Mohler; see his papers with Rink [208, 224]. In the Russian control literature Buyakas [46], Barbašin [21, 22] and others wrote about them. Mathematical work on (1.22) began with Kučera [167–169] and gained impetus from the efforts of Brockett [34–36]. Among the surveys of the earlier literature on bilinear control are Bruni et al. [43], Mohler [206], [207, Vol. II], and Elliott [83] as well as the proceedings of seminars organized by Mohler and Ruberti [204, 209, 227].

Example 1.1. Bilinear control systems are in many instances intentionally designed by engineers. For instance, an engineer designing an industrial process might begin with the simple mathematical model $\dot{x} = Ax + bu$ with one input and one observed output $y = c^\top x$. Then a simple control system design might be $u = v(t)y$, the action of a proportional valve v with $0 \le v(t) \le 1$. Thus the control system will be bilinear, $\dot{x} = Ax + v(t)Bx$ with $B = bc^\top$ and $\Omega := [0, 1]$. The valve setting $v(t)$ might be adjusted by a technician or by automatic computation. For more about bilinear systems with $B = bc^\top$, see Exercise 1.12 and Section 6.4. △

Mathematical models arise in hierarchies of approximation. One common nonlinear model that will be seen again in later chapters is (1.21), which often can be understood only qualitatively and by computer simulation, so simple approximations are needed for design and analysis. Near a state ξ

[12] Feedback control is also called closed-loop control.

[13] Here $-\Omega := \{-u \mid u \in \Omega\}$.

with $f(\xi) = a$, the control system (1.21) can be linearized in the following way. Let

$$A = \frac{\partial f}{\partial x}(\xi),\ b_i = g_i(\xi),\ B_i = \frac{\partial g_i}{\partial x}(\xi).\ \text{Then} \tag{1.24}$$

$$\dot{x} = Ax + a + \sum_i^m u_i(t)(B_i x + b_i) \tag{1.25}$$

to first order in x. A control system described by (1.25) can be called an inhomogeneous[14] bilinear system; (1.25) has affine vector fields ($Bx + b$), so I prefer to call it a biaffine control system.[15] Biaffine control systems will be the topic of Section 3.8 in Chapter 3; in this chapter, see Exercise 1.15. In this book, *bilinear control systems* (1.22) have $a = 0$ and $b_i = 0$ and are called *homogeneous* when the distinction must be made.

Bilinear and biaffine systems as intentional design elements (Example 1.1) or as approximations via (1.24) are found in many areas of engineering and science:

- chemical engineering — valves, heaters, catalysts;
- mechanical engineering — automobile controls and transmissions;
- electrical engineering — frequency modulation, switches, voltage converters;
- physics — controlled Schrödinger equations, spin dynamics;
- biology — neuron dendrites, enzyme kinetics, compartmental models.

For descriptions and literature citations of several such applications see Chapter 6. Some of them use discrete-time bilinear control systems[16] described by difference equations on $\mathbb{F}^n$

$$x(t+1) = Ax(t) + \sum_1^m u_i(t)B_i x(t),\ u(\cdot) \in \Omega,\ t \in \mathbb{Z}_+. \tag{1.26}$$

Again a feedback control $u = \phi(x)$ may be considered by the designer. If ϕ is continuous, a nonlinear difference equation of this type will have a unique solution, but for a bad choice of ϕ the solution may grow exponentially or faster; look at $\overset{+}{x} = -x\phi(x)$ with $\phi(x) := x$ and $x(0) = 2$.

Section 1.8 is a sidebar about some discrete-time systems that arise as methods of approximately computing the trajectories of (1.22). Only one of them (Euler's) is bilinear. Controllability and stabilizability theories for

---

[14] Rink and Mohler [224] used the terms "bilinear system" and "bilinear control process" for systems like (1.25). The adjective "biaffine" was used by Sontag, Tarn and others, reserving "bilinear" for (1.22).

[15] Biaffine systems can be called *variable structure systems* [204], an inclusive term for control systems affine in x (see Exercise 1.17) that includes switching systems with sliding modes; see Example 6.8 and Utkin [281].

[16] For early papers on discrete-time bilinear systems, see Mohler and Ruberti [204].

discrete-time bilinear systems are sufficiently different from the continuous case to warrant Chapter 4; their observability theory is discussed in parallel with the continuous-time version in Chapter 5.

Given an indexed family of matrices $\mathcal{F} = \{B_j |\, j \in 1\,..\,m\}$ (where m is in most applications a small integer), we define a switching system on $\mathbb{R}^n_*$ as

$$\dot{x} = B_{j(t)}x,\ t \in \mathbb{R}_+,\ j(t) \in 1\,..\,m. \tag{1.27}$$

The control input $j(t)$ for a switching system is the assignment of one of the indices at each t.[17] The switching system described in (1.27) is a special case of (1.26) with no drift term and can be written

$$\dot{x} = \sum_{i=1}^{m} u_i B_i \ \text{ where } \ u_i(t) = \delta_{i,j(t)},\ j(t) \in 1\,..\,m;$$

so (1.27) is a bilinear control system. It is also a linear dynamical polysystem as defined in Section B.3.4.

Examples of switching systems appear not only in electrical engineering but also in many less obvious contexts. For instance, the B_i can be matrix representations of a group of rotations or permutations.

Remark 1.3. To avoid confusion, it is important to mention the discrete-time control systems that are described not by (1.26) but instead by input-to-output mappings of the form $y = f(u_1, u_2)$ where u_1, u_2, y are sequences and f is a bilinear function. In the 1970s, these were called "bilinear systems" by Kalman and others. The literature of this topic, for instance Fornasini and Marchesini [94, 95] and subsequent papers by those authors, became part of the extensive and mathematically interesting theories of dynamical systems whose time-set is $\mathbb{Z}^2$ (called *2-D systems*); such systems are discussed, treating the time-set as a discrete free semigroup, in Ball et al. [18]. △

1.5 Transition Matrices

Consider the controlled dynamics on $\mathbb{R}^n$

$$\dot{x} = \left(A + \sum_{i=1}^{m} u_i(t)B_i\right)x,\ x(0) = \xi;\ u \in \mathcal{PK},\ u(\cdot) \in \Omega. \tag{1.28}$$

As a control system, (1.28) is time-invariant. However, once given an input history $\mathfrak{U}_T := \{u(t) |\, 0 \le t \le T\}$, (1.28) can be treated as a linear time-variant

[17] In the continuous-time case of (1.27), there must be some restriction on how often j can change, which leads to some technical issues if $j(t)$ is a random process as in Section 8.3.

(but piecewise constant) vector differential equation. To obtain transition mappings $x(t) = \Theta_t(\xi; u)$ for any ξ and u, we will show in Section 1.5.1 that on the linear space $L = \mathbb{R}^{n\times n}$ the matrix control system

$$\dot{X} = AX + \sum_{i=1}^{m} u_i(t)B_iX,\ X(0) = I,\ u(\cdot) \in \Omega \tag{1.29}$$

with piecewise constant ($\mathcal{PK}$) controls has a unique solution $X(t; u)$, called the transition matrix for (1.28); $\Theta_t(\xi; u) = X(t; u)\xi$.

1.5.1 Construction Methods

The most obvious way to construct transition matrices for (1.29) is to represent $X(t; u)$ as a product of exponential matrices.

A $u \in \mathcal{PK}$ defined on a finite interval $[0, T]$ has a finite number N of intervals of constancy that partition $[0, T]$. What value u has at the endpoints of the N intervals is immaterial. Given u, and any time $t \in [0, t]$ there is an integer $k \le N$ depending on t such that

$$u(s) = \begin{cases} u(0), & 0 = \tau_0 \le s < \tau_1, \dots, \\ u(\tau_{i-1}), & \tau_{i-1} \le s < \tau_i, \\ u(\tau_k), & \tau_k \le= t; \end{cases}$$

let $\sigma_i = \tau_i - \tau_{i-1}$. The descending-order product

$$X(t; u) := \exp\left((t - \tau_k)\left(A + \sum_{j=1}^{m} u_j(\tau_k)B_j\right)\right)\prod_{i=k}^{1}\exp\left(\sigma_i\left(A + \sum_{j=1}^{m} u_j(\tau_{i-1})B_j\right)\right) \tag{1.30}$$

is the desired transition matrix. It is piecewise differentiable of all orders, and estimates like those used in Proposition 1.1 show it is unique.

The product in (1.30) leads to a way of constructing solutions of (1.29) with piecewise continuous controls. Given a bounded control $u \in \mathcal{PC}[0, T]$, approximate u with a sequence $\{u^{(1)}, u^{(2)}, \dots\}$ in $\mathcal{PK}[0, T]$ such that

$$\lim_{i\to\infty} \int_0^T \|u^{(i)}(t) - u(t)\|\, dt = 0.$$

Let a sequence of matrix functions $X(u^{(i)}; t)$ be defined by (1.30). It converges uniformly on $[0, T]$ to a limit called a multiplicative integral or product integral,

$$X(u^{(N)}; t) \xrightarrow[N\to\infty]{} X(t; u),$$

that is, a continuously differentiable solution of (1.29). See Dollard and Friedman [74] for the construction of product integral solutions to $\dot{X} = A(t)X$ for $A(t)$ piecewise constant (a *step function*) or locally integrable.

Another useful way to construct transition matrices is to pose the initial-value problem as an integral equation, then to solve it iteratively. Here is an example; assume that u is integrable and essentially bounded on $[0, T]$. If

$$\dot{X} = AX + u(t)BX,\ X(0) = I;\ \text{ then} \tag{1.31}$$

$$X(t; u) = e^{tA} + \int_0^t u(t_1)e^{(t-t_1)A}BX(t_1; u)\,dt_1,\ \text{so formally} \tag{1.32}$$

$$X(t; u) = e^{tA} + \int_0^t u(t_1)e^{(t-t_1)A}Be^{t_1A}\,dt_1 +$$
$$\int_0^t\int_0^{t_1} u(t_1)u(t_2)e^{(t-t_1)A}Be^{(t_1-t_2)A}Be^{t_2A}\,dt_2\,dt_1 + \cdots . \tag{1.33}$$

This Peano–Baker series is a generalization of (1.6), and the same kind of argument shows that (1.33) converges uniformly on to $[0, T]$ [249, Ch. 11] to a differentiable matrix function that satisfies (1.31). The Peano–Baker series will be useful for constructing Volterra series in Section 5.6. For the construction of formal Chen–Fliess series for $X(t; u)$ from (1.32), see Section 8.2.

Example 1.2. If $AB = BA$, then $e^{A+B} = e^Ae^B = e^Be^A$ and the solution of (1.31) satisfies $X(t; u) = \exp(\int_0^t (A + u(s)B)\,ds)$. However, suppose $AB \neq BA$. Let $u(t) = 1$ on $[0, \tau/1)$, $u(t) = -1$ on $[\tau/2, \tau)$,

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix};\ X(\tau; u) = e^{\tau(A-B)/2}e^{\tau(A+B)/2} = \begin{bmatrix} 1 & \tau \\ \tau & 1+\tau^2 \end{bmatrix},\ \text{ but}$$
$$\exp\left(\int_0^\tau (A + u(s)B)\,ds\right) = \begin{bmatrix} \cosh(\tau) & \sinh(\tau) \\ \sinh(\tau) & \cosh(\tau) \end{bmatrix}. \qquad \triangle$$

Example 1.3 (Polynomial inputs). What follows is a classical result for time-variant linear differential equations in a form appropriate for control systems. Choose a polynomial input $u(t) = p_0 + p_1t + \cdots + p_kt^k$ and suppose that the matrix formal power series

$$X(t) := \sum_{i=0}^{\infty} t^iF_i \text{ satisfies } \dot{X} = (A + u(t)B)X,\ X(0) = I;\ \text{then}$$

$$\dot{X}(t) = \sum_{i=0}^{\infty} t^i(i+1)F_{i+1} = \left(A + B\sum_{i=0}^{k} p_it^i\right)\left(\sum_{i=0}^{\infty} t^iF_i\right);$$

evaluating the product on the right and equating the coefficients of t^i

$$(i+1)F_{i+1} = AF_i + B\sum_{j=0}^{i\wedge k} F_j p_{i-j},\ F_0 = I, \tag{1.34}$$

where $i \wedge k := \min(i,k)$; all the F_i are determined.

To show that the series $X(t)$ has a positive radius of convergence, we only need to show that $\|F_i\|$ is uniformly bounded. Let $p_i := \|F_i\|$, $\alpha := \|A\|$, $\beta := \|B\|$, and $\mu_i := p_i$; then $\{p_i, i \in \mathbb{Z}_+\}$ satisfies the system of linear time-variant inequalities

$$(i+1)p_{i+1} \le \alpha p_i + \beta\sum_{j=0}^{i\wedge k} p_j \mu_{i-j},\ \text{with } p_0 = 1. \tag{1.35}$$

Since eventually $i+1$ is larger than $(k+1)\max\{\alpha, \beta\mu_1, \dots, \beta\mu_k\}$, there exists a uniform bound for the p_i. To better see how the F_i behave as $i \to \infty$, as an exercise consider the commutative case $AB = BA$ and find the coefficients F_i of the power series for $\exp(\int_0^t (A + u(s)B)\,ds)$. △

Problem 1.1. (a) In Example 1.3, the sequence of inequalities (1.35) suggests that there exists a number C such that $p_i \le C/\Gamma(i/k)$, where $\Gamma(\cdot)$ is the usual gamma-function, and that the series $X(t)$ converges on $\mathbb{R}$.

(b) If u is given by a convergent power series, show that $X(t)$ converges for sufficiently small $|t|$. △

1.5.2 Semigroups

A semigroup is a set **S** equipped with an associative binary operation (called composition or multiplication) $\mathbf{S} \times \mathbf{S} \to \mathbf{S} : (a,b) \mapsto ab$ defined for all elements $a, b, \dots \in \mathbf{S}$. If **S** has an element ι (the identity) such that $\iota a = a\iota$, it should be called a semigroup with identity or a monoid; it will be called a semigroup here. If $ba = \iota = ab$ then $a^{-1} := b$ is the inverse of a. If every element of **S** has an inverse in **S**, then **S** is a group; see Section B.4 for more about groups.

A semigroup homomorphism is a mapping $h : \mathbf{S}_1 \to \mathbf{S}_2$ such that $h(\iota_1) = \iota_2$ and for $a, b \in \mathbf{S}_1$, $h(ab) = h(a)h(b)$. An example is as follows:

$$h_A : \mathbb{R}_+ \to \exp(A\mathbb{R}_+);\ h_A(t) = \exp(tA),\ \iota_1 = 0,\ h_A(0) = I.$$

Define $\mathbf{SS} := \{XY | X \in S,\ Y \in S\}$. A subsemigroup of a group **G** is a subset $\mathbf{S} \subset \mathbf{G}$ that satisfies $\iota \in \mathbf{S}$ and $\mathbf{SS} = \mathbf{S}$.

Definition 1.5. A matrix semigroup **S** is a subset of $\mathbb{F}^{n\times n}$ that is closed under matrix multiplication and contains the identity matrix I_n. The elements of **S** whose inverses belong to **S** are called units. △

Example 1.4. There are many more or less obvious examples of semigroups of matrices, for instance:

(a) $\mathbb{R}^{n\times n}$, the set of all real matrices.
(b) The upper triangular matrices $\{A \in \mathbb{R}^{n\times n} | \, a_{i,j} = 0, i < j\}$.
(c) The nonnegative matrices $\{A \in \mathbb{R}^{n\times n} | \, a_{i,j} \geq 0\}$; these are the matrices that map the (closed) positive orthant $\mathbb{R}^n_+$ into itself.
(d) The doubly stochastic matrices, which are the subset of (e) for which all row-sums and column-sums equal one.

The reader should take a few minutes to verify that these are semigroups. For foundational research on subsemigroups of Lie groups, see Hilgert et al. [127, Ch. V]. △

A concatenation semigroup $\mathfrak{S}_\star$ can be defined for inputs in $\mathcal{U} = \mathcal{LI}, \mathcal{PK}$, or $\mathcal{PC}$ of finite duration whose values lie in Ω. $\mathfrak{S}_\star$ is a set of equivalence classes whose representatives are pairs $\bar{u} := (u, [0, T_u])$: a function u with values in Ω and its interval of support $[0, T_u]$. Two inputs $\bar{u}, \bar{v}$ are equivalent, written $\bar{u} \sim \bar{v}$, if $T_u = T_v$ and $u(t) = v(t)$ for almost all $t \in [0, T_u]$. The semigroup operation $\bar{\star}$ for $\mathfrak{S}_\star$ is given by

$$(\bar{u}, \bar{v}) \mapsto \bar{u} \bar{\star} \bar{v} = (u \star v, [0, T_u + T_v]),$$

where $\star$ is defined as follows. Given two scalar or vector inputs, say $(u, [0, \tau])$ and $(v, [0, \sigma])$, their concatenation $u \star v$ is defined on $[0, \sigma + \tau]$ by

$$u \star v(t) = \begin{cases} v(t), & t \in [0, \tau], \\ u(t - \tau), & t \in [\tau, \sigma + \tau], \end{cases} \tag{1.36}$$

since its value at the point $t = \tau$ can be assigned arbitrarily. The identity element in $\mathfrak{S}_\star$ is $\{\bar{u} | \, \bar{u} \sim (0, 0)\}$.

Given $\bar{u} = (u, [0, T])$ and a time $\tau < T$, the control u can be split into the concatenation of u', the restriction of u to $[0, \tau]$, and u'', the restriction of S$-\tau u$ to $[0, T - \tau]$:

$$\begin{gathered} u = u'' \star u' \quad \text{where} \\ u'(t) = u(t), \ 0 \leq t \leq \tau; \ u''(t) = u(t + \tau), \ 0 \leq t \leq T - \tau. \end{gathered} \tag{1.37}$$

The connection between $\mathfrak{S}_\star$ and (1.29) is best seen in an informal, but traditional, way using (1.37). In computing a trajectory of (1.29) for control u on $[0, T]$ one can halt at time τ, thus truncating u to u' and obtaining an intermediate transition matrix $Y = X(\tau; u')$. Restarting the computation using u'' with initial condition $X(0) = Y$, on $[\tau, T]$, we find $X(t; u) = X(t - \tau; u'')Y$. One obtains, this way, the transition matrix trajectory

$$X(t; u) = \begin{cases} X(t; u'), \ 0 \leq t \leq \tau, \\ X(t; u'' \star u') = X(t - \tau; u'')X(\tau; u'), \ \tau < t \leq T. \end{cases} \tag{1.38}$$

With the above setup, the relation $X(t;u) = X(t-\tau;u'')X(\tau;u')$ is called the semigroup property (or the composition property) of transition matrices.

Definition 1.6. $\mathbf{S}_\Omega$ is the smallest semigroup containing the set

$$\left\{ \exp\left(\tau A + \sum_{i=1}^{m} u_i \tau B_i \right) \middle| \tau \geq 0,\ u \in \Omega \right\}.$$

If $\Omega = \mathbb{R}^m$ the subscript of $\mathbf{S}$ is omitted. △

A representation of $\mathbf{S}_\Omega$ on $\mathbb{R}^n$ is a semigroup homomorphism $\rho : \mathbf{S}_\Omega \to \mathbb{R}^{n\times n}$ such that $\rho(\iota) = I_n$. By its definition and (1.30), the matrix semigroup $\mathbf{S}_\Omega$ contains all the transition matrices $X(t;u)$ satisfying (1.29) with $u \in \mathcal{PK}$, and each element of $\mathbf{S}_\Omega$ is a transition matrix. Using (1.38), we see that the mapping $X : \mathfrak{S}_\star \to \mathbf{S}_\Omega$ is a representation of $\mathfrak{S}_\star$ on $\mathbb{R}^n$. For topologies and continuity properties of these semigroup representations, see Section 8.1.

Proposition 1.5. *Any transition matrix $X(t;u)$ for* (1.28) *satisfies*

$$\det(X(t;u)) = \exp\left(\int_0^t \left(\mathrm{tr}(A) + \sum_{j=1}^{m} u_j(s)\, \mathrm{tr}(B_j) \right) ds \right) > 0. \tag{1.39}$$

Proof. Using (1.9) and $\det(X^k) = (\det(X))^k$, this generalization (1.39) of Abel's relation holds for u in $\mathcal{PK}$, $\mathcal{LI}$, and $\mathcal{PC}$. □

As a corollary, every transition matrix Y has an inverse Y^{-1}; but Y is not necessarily a unit in $\mathbf{S}_\Omega$. That may happen if $A \neq 0$ or if the value of u is constrained; Y^{-1} may not be a transition matrix for (1.28). This can be seen even in the simple system $\dot{x}_1 = u_1 x_1$ with the constraint $u_1 \geq 0$.

Proposition 1.6. *Given a symmetric bilinear system with $u \in \mathcal{PK}$, each of its transition matrices $X(\tau;u)$ is a unit; $(X(\tau;u))^{-1}$ is again a transition matrix for* (1.23).

Proof. This, like the composition property, is easy using (1.30) and remains true for piecewise continuous and other locally integrable inputs. Construct the input

$$u^-(t) = -u(\tau - t),\ 0 \leq t \leq \tau, \tag{1.40}$$

which undoes the effect of the given input history. If $X(0,u^-) = I$ and

$$\dot{X} = (u_1^- B_1 + \ldots + u_k^- B_k)X, \text{ then}$$
$$X(\tau;u^-)X(\tau;u) = I; \text{ similarly}$$
$$X(2\tau u \star u^-) = I = X(2\tau;u^- \star u).$$

□

1.5.3 Matrix Groups

Definition 1.7. A matrix group[18] is a set of matrices $\mathbf{G} \subset \mathbb{R}^{n\times n}$ with the properties that (a) $I \in \mathbf{G}$ and (b) if $X \in \mathbf{G}$ and $Y \in \mathbf{G}$ then $XY^{-1} \in \mathbf{G}$. For matrix multiplication $X(YZ) = (XY)Z$, so $\mathbf{G}$ is a group. △

There are many different kinds of matrix groups; here are some well-known examples.

1. The group $\mathrm{GL}(n, \mathbb{R})$ of all invertible real $n \times n$ matrices is called the real general linear group. $\mathrm{GL}(n, \mathbb{R})$ has two connected components, separated by the hypersurface $\Delta_0^n := \{X|\ \det(X) = 0\}$. The component containing I (identity component) is

$$\mathrm{GL}^+(n, \mathbb{R}) := \{X \in \mathrm{GL}(n, \mathbb{R})|\ \det(X) > 0\}.$$

   From (1.9) we see that for (1.29), $X(t; u) \in \mathrm{GL}^+(n, \mathbb{R})$ and the transition matrix semigroup $\mathbf{S}_\Omega$ is a subsemigroup of $\mathrm{GL}^+(n, \mathbb{R})$.
2. $\mathbf{G} \subset \mathrm{GL}(n, \mathbb{F})$ is called finite if its cardinality $\#(\mathbf{G})$ is finite, and $\#(\mathbf{G})$ is called its order. For any $A \in \mathrm{GL}(n, \mathbb{F})$ for which there exists $k > 0$ such that $A^k = I$, $\{I, A, \ldots, A^{k-1}\}$ is called a cyclic group of order k.
3. The discrete group $\mathrm{SL}(n, \mathbb{Z}) := \{A \subset \mathrm{GL}(n, \mathbb{R})|\ a_{i,j} \in \mathbb{Z}, \det(A) = 1\}$ is infinite if $n > 1$, since it contains the infinite subset

$$\left\{\begin{bmatrix} j+1 & j+2 \\ j & j+1 \end{bmatrix} \middle|\ j \in \mathbb{Z}\right\}.$$

4. For any $A \in \mathbb{R}^{n\times n}$

$$\mathrm{e}^{rA}\mathrm{e}^{tA} = \mathrm{e}^{(r+t)A} \text{ and } (\mathrm{e}^{tA})^{-1} = \mathrm{e}^{-tA}, \text{ so } \mathbb{R} \to \mathrm{GL}^+(n, \mathbb{R}) :\ t \mapsto \mathrm{e}^{tA}$$

   is a group homomorphism (Section B.8); the image is called a one–parameter subgroup of $\mathrm{GL}^+(n, \mathbb{R})$.

Rather than using the semigroup notation $\mathbf{S}$, the set of transition matrices for symmetric systems will be given its own symbol.

Proposition 1.7. *For a given* symmetric *system* (1.23) *with piecewise constant controls as in* (1.30)*, the set of all transition matrices is given by the set of finite products*

$$\boldsymbol{\Phi} := \left\{\prod_{j=k}^{1} \exp\Big(\sigma_j \sum_{i=1}^{m} u_i(\tau_j) B_i\Big) \,\middle|\, \sigma_j \geq 0,\ u(\cdot) \in \Omega, k \in \mathbb{N}\right\}. \tag{1.41}$$

$\boldsymbol{\Phi}$ is a connected matrix subgroup of $\mathrm{GL}^+(n, \mathbb{R})$.

---

[18] See Section B.4.

Proof. Connected in this context means path-connected, as that term is used in general topology. By definition $I \in \boldsymbol{\Phi}$; by (1.9), $\boldsymbol{\Phi} \subset \mathrm{GL}^+(n, \mathbb{R})$. Any two points in $\boldsymbol{\Phi}$ can be connected to I by a finite number of arcs in $\boldsymbol{\Phi}$ of the form $\exp(tB_i)$, $i \in 1 \dots m$, $0 \le t \le T$. That $\boldsymbol{\Phi}$ is a group follows from the composition property (1.38) and Proposition 1.6; from that it follows that any two points in $\boldsymbol{\Phi}$ are path-connected. □

Example 1.5. Consider, for $\Omega = \mathbb{R}$, the semigroup **S** of transition matrices for the system $\dot{x} = (J + uI)x$ on $\mathbb{R}^2$ where $J = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. If for $t > 0$ and given u

$$X(t; u) = e^{\int_0^t u(s)\,ds} \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}, \text{ then}$$
$$\text{if } k \in \mathbb{N} \text{ and } 2k\pi > t, \quad (X(t; u))^{-1} = X(-u; 2k\pi - t).$$

Therefore **S** is a group, isomorphic to the group $\mathbb{C}_*$ of nonzero complex numbers.[19] △

1.6 Controllability

A control system is to a dynamical system as a play is to a monolog: instead of a single trajectory through ξ one has a family of them. One may look for the subset of state space that they cover or for trajectories that are optimal in some way. Such ideas led to the concept of state-to-state controllability, which first arose when the calculus of variations[20] was applied to optimal control. Many of the topics of this book invoke controllability as a hypothesis or as a conclusion; its general definition (Definition 1.9) will be prefaced by a well-understood special case: controllable linear systems.

1.6.1 Controllability of Linear Systems

The linear control system (1.18) is said to be controllable on $\mathbb{R}^n$ if for each pair of states $\{\xi, \zeta\}$ there exists a time $\tau > 0$ and a locally integrable control $u(\cdot) \in \Omega \subset \mathbb{R}^m$ for which

$$\zeta = e^{\tau A}\xi + \int_0^\tau e^{(\tau - s)A} B u(s)\,ds. \tag{1.42}$$

---

[19] See Example 2.4 in the next chapter for the $\boldsymbol{\alpha}$ representation of $\mathbb{C}$.

[20] The book of Hermann [122] on calculus of variations stimulated much early work on the geometry of controllability.

If the control values are unconstrained ($\Omega = \mathbb{R}^m$), the necessary and sufficient Kalman controllability criterion for controllability of (1.18) is

$$\operatorname{rank} \mathbf{R} = n, \text{ where } \mathbf{R}(A;B) := \left[B\ AB\ \dots\ A^{n-1}B\right] \in R^{n\times mn}. \tag{1.43}$$

$\mathbf{R}(A;B)$ is called the controllability matrix. If (1.43) holds, then $\{A,B\}$ is called a controllable pair. If there exists some vector b for which $\{A,b\}$ is a controllable pair, A is called cyclic. For more about linear systems see Exercises 1.11–1.14. The criterion (1.43) has a dual, the observability criterion (5.2) for linear dynamical systems at the beginning of Chapter 5; for a treatment of the duality of controllability and observability for linear control systems, see Sontag [249, Ch. 5].

Definition 1.8. A subset K of any linear space is said to be a convex set if for all points $x \in K$, $y \in K$ and for all $\mu \in [0,1]$, we have $\mu x + (1-\mu)y \in K$.

△

Often the control is constrained to lie in a closed convex set $\Omega \subset \mathbb{R}^m$. For linear systems (1.18), if Ω is convex then

$$\left\{\int_0^\tau e^{(\tau-s)A}Bu(s)\,ds,\ u(\cdot)\in\Omega\right\} \text{ is convex.} \tag{1.44}$$

1.6.2 *Controllability in General*

It is convenient to define controllability concepts for control systems more general than we need just now. Without worrying at all about the technical details, consider a nonlinear control system on a connected region $U \subset \mathbb{R}^n$,

$$\dot{x} = f(x,u),\ u \in \mathcal{LI},\ t \in \mathbb{R}_+,\ u(\cdot)\in\Omega. \tag{1.45}$$

Make technical assumptions on f and U such that there exists for each u and for each $\xi \in U$ a controlled trajectory $f_{t;u}(\xi)$ in U with $x(0) = f_{0;u}(\xi) = \xi$ that satisfies (1.45) almost everywhere.

Definition 1.9 (Controllability). The system (1.45) is said to be controllable on U if given any two states $\xi, \zeta \in U$ there exists $T \geq 0$ and control u for which $f_{t;u}(\xi) \in U$ for $t \in (0,T)$, and $f_{T;u}(\xi) = \zeta$. We say that with this control system the target ζ is reachable from ξ using u. △

Definition 1.10 (Attainable sets).[21] For a given instance of (1.45), for initial state $\xi \in U$:

(i) The t-attainable set is $\mathcal{A}_t(\xi) := \{f_{t;u}(\xi) \,|\, u(\cdot)\in\Omega\}$.

[21] Jurdjevic [147] and others call them reachable sets.

(ii) The set of all states attainable by time T is $\mathcal{A}^T(\xi) := \bigcup_{0<t\leq T} \mathcal{A}_t(\xi)$.
(iii) The attainable set is $\mathcal{A}(\xi) := \bigcup_{t\geq 0} \mathcal{A}_t(\xi)$. △

We will occasionally need the following notations. Given a set $U \subset \mathbb{R}^k$, its complement is denoted by $\mathbb{R}^k \setminus U$, its closure by $\overline{U}$, its interior by $\overset{\circ}{U}$, and its boundary by $\partial U = \overline{U} \setminus \overset{\circ}{U}$. For instance, it can happen that $\mathcal{A}(\xi)$ includes part of $\partial\mathcal{A}(\xi)$, so may be neither open nor closed in $\mathbb{R}^n$.

Definition 1.11. i) If $\overset{\circ}{\mathcal{A}}(\xi) \neq \emptyset$ one says that the accessibility condition is satisfied at ξ. If the accessibility condition is satisfied for all initial states $\xi \in U$, we say that (1.45) has the accessibility property on U.

ii) If, for each $t > 0$, $\overset{\circ}{\mathcal{A}}_t(\xi) \neq \emptyset$ the control system has the property of strong accessibility from ξ; if that is true for all $\xi \neq 0$, then (1.28) is said to satisfy the strong accessibility condition on U. △

The following lemma, from control-theory folklore, will be useful later. Besides (1.45), introduce its discrete-time version

$$\overset{+}{x} = F(x, u),\ t \in \mathbb{Z}_+,\ u(\cdot) \in \Omega, \tag{1.46}$$

a corresponding notion of controlled trajectory $\boldsymbol{F}_{t;u}(\xi)$ and the same definition of controllability.

Lemma 1.1 (Continuation). *The system* (1.45) [or (1.46)] *is controllable on a connected submanifold $U \subset \mathbb{R}^n$ if for every initial state ξ there exists u such that a neighborhood $N(\xi) \subset U$ is attainable by a controlled trajectory lying in U.*

Proof. Given any two states $\xi, \zeta \in U$, we need to construct on U a finite set of points $p_0 = \xi, \ldots, p_k = \zeta$ such that a neighborhood in U of p_j is attainable from p_{j-1}, $j \in 1, \ldots, k$. First construct a continuous curve $\gamma : [0, 1] \to U$ of finite length such that $\gamma(0) = \xi$, $\gamma(1) = \zeta$. For each point $p = \gamma(\tau)$, there exists an attainable neighborhood $N(p) \subset U$. The collection of neighborhoods $\bigcup_x N(x)$ covers the compact set $\{\gamma(t), t \in [0, 1]\}$, so we can find a finite subcover $N_1, \ldots, N_k$ where $p_1 \in N(\xi)$ and the k neighborhoods are ordered along γ. From ξ, choose a control that reaches $p_1 \in N_1 \cap N_2$, and so forth; $\zeta \in N_k$ can be reached from p_{k-1}. □

1.6.3 Controllability: Bilinear Systems

For homogeneous bilinear systems (1.22), the origin in $\mathbb{R}^n$ is an equilibrium state for all inputs; for that reason in Definition 1.10, it is customary to use as the region U the punctured space $\mathbb{R}^n_* := \mathbb{R}^n \setminus \{0\}$. The controlled trajectories are $x(t) = X(t; u)\xi$; the attainable set from ξ is

$$\mathcal{A}(\xi) := \{X(t;u)\xi \,|\, t \in \mathbb{R}_+,\ u(t) \in \Omega\};$$

if $\mathcal{A}(\xi) = \mathbb{R}^n_*$ for all $\xi \in \mathbb{R}^n_*$ then (1.28) is controllable on $\mathbb{R}^n_*$.

In one dimension, $\mathbb{R}^1_*$ is the union of two open disjoint half-lines; therefore, no continuous-time system[22] can be controllable on $\mathbb{R}^1_*$, so our default assumption about $\mathbb{R}^n_*$ is that $n \geq 2$. The punctured plane $\mathbb{R}^2_*$ is connected but not simply connected. For $n \geq 3$, $\mathbb{R}^n_*$ is simply connected.

Example 1.6. Here is a symmetric bilinear system that is uncontrollable on $\mathbb{R}^2_*$.

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}u_1x_1\\ u_2x_2\end{bmatrix},\ x(0) = \xi;\ x(t) = \begin{bmatrix}\exp(\int_0^t u_1(s)\,ds)\xi_1\\ \exp(\int_0^t u_2(s)\,ds)\xi_2\end{bmatrix}.$$

Two states ξ, x can be connected by a trajectory of this system if and only if (i) ξ and x lie in the same open quadrant of the plane or (ii) ξ and x lie in the same component of the set $\{x \,|\, x_1x_2 = 0\}$. △

Example 1.7. For the two-dimensional system

$$\dot{x} = (I + u(t)J)x,\ I = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\ J = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} \tag{1.47}$$

for $t > 0$ the t-attainable set is $\mathcal{A}_t(\xi) = \{x \,|\, \|x\| = \|\xi\|e^t\}$, a circle, so (1.47) does not have the strong accessibility property anywhere. The attainable set is $\mathcal{A}(\xi) = \{x \,|\, \|x\| > \|\xi\|\} \bigcup \{\xi\}$, so (1.47) is not controllable on $\mathbb{R}^2_*$. However, $\mathcal{A}(x)$ has open interior, although it is neither open nor closed, so the accessibility property is satisfied. In the $\mathbb{R}^2 \times \mathbb{R}_+$ space-time, $\{(\mathcal{A}_t(\xi), t) \,|\, t > 0\}$ is a truncated circular cone. △

Example 1.8. Now reverse the roles of I, J in the last example, giving $\dot{x} = (J + uI)x$ as in Example 1.5; in polar coordinates this system becomes $\dot{\rho} = u\rho$, $\dot{\theta} = 1$. At any time $T > 0$, the radius ρ can have any positive value, but $\theta(T) = 2\pi T + \theta_0$. This system has the accessibility property (not strong accessibility) and *it is controllable*, although any desired final angle can be reached only once per second. In space-time the trajectories of this system are $\{(\int^t u, \theta_0 + 2\pi t, t) \,|\, t \geq 0\}$ and thus live on the positive half of a helicoid surface.[23] △

1.6.4 *Transitivity*

A matrix semigroup $\mathbf{S} \subset \mathbb{R}^{n\times n}$ is said to to be transitive on $\mathbb{R}^n_*$ if for every pair of states $\{x, y\}$ there exists at least one matrix $A \in \mathbf{S}$ such that $Ax = y$.

[22] The discrete-time system $x(t + 1) = u(t)x(t)$ is controllable on $\mathbb{R}^1_*$.

[23] In Example 1.8, the helicoid surface $\{(r, 2\pi t, t) \,|\, t \in \mathbb{R}, r > 0\}$ covers $\mathbb{R}^2_*$ exactly the way the Riemann surface for $\log(z)$ covers $\mathbb{C}_*$, with fundamental group $\mathbb{Z}$.

If a given bilinear system is controllable, its semigroup of transition matrices $\mathbf{S}_\Omega$ is transitive on $\mathbb{R}^n_*$; a parallel result for the transition matrices of a discrete-time system will be discussed in Chapter 4. For symmetric systems, controllability is just the transitivity of the matrix group $\boldsymbol{\Phi}$ on $\mathbb{R}^n_*$. For systems with drift, often the way to prove controllability is to show that $\mathbf{S}_\Omega$ is actually a transitive group.

Example 1.9. An example of a semigroup (without identity) of matrices transitive on $\mathbb{R}^n_*$ is the set of matrices of rank k,

$$\mathbf{S}^n_k := \{A \in \mathbb{R}^{n\times n} | \operatorname{rank} A = k\}; \text{ since } \mathbf{S}^n_k\mathbf{S}^n_k = \mathbf{S}^n_k, \tag{1.48}$$

the $\mathbf{S}^n_k$ are semigroups but for $k < n$ have no identity element. To show the transitivity of $\mathbf{S}^n_1$, given $x, y \in \mathbb{R}^n_*$ let $A = yx^\top/(x^\top x)$; then $Ax = y$. Since for $k > 1$ these semigroups satisfy $\mathbf{S}^n_k \supset \mathbf{S}^n_{k-1}$, they are all transitive on $\mathbb{R}^n_*$, including the biggest, which is $\mathbf{S}^n_n = \mathrm{GL}(n, \mathbb{R})$. △

Example 1.10. Transitivity has many contexts. For $A, B \in \mathrm{GL}(n, \mathbb{R})$, there exists $X \in \mathrm{GL}(n, \mathbb{R})$ such that $XA = B$. The columns of $A = (a_1, \dots, a_n)$ are a linearly independent n-tuple of vectors. Such an n-tuple A is called an n-frame on $\mathbb{R}^n$; so $\mathrm{GL}(n, \mathbb{R})$ is transitive on n-frames. For n-frames on differentiable n-manifolds, see Conlon [64, Sec. 5.5]. △

1.7 Stability: Nonlinear Dynamics

In succeeding chapters, stability questions for *nonlinear* dynamical systems will arise, especially in the context of state-dependent controls $u = \phi(x)$. The best-known tool for attacking such questions is Lyapunov's direct method. In this section, our dynamical system has state space $\mathbb{R}^n$ with transition mapping generated by a differential equation $\dot{x} = f(x)$, assuming that f satisfies conditions guaranteeing that for every $\xi \in \mathbb{R}^n$ there exists a unique trajectory in $\mathbb{R}^n$

$$x(t) = f_t(\xi),\ 0 \le t \le \infty.$$

First locate the equilibrium states x_e that satisfy $f(x_e) = 0$; the trajectory $f_t(x_e) = x_e$ is a constant point. To investigate an isolated[24] equilibrium state x_e, it is helpful to change the coordinates so that $x_e = 0$, the origin; that will now be our assumption. .

The state 0 is called stable if there exists a neighborhood $\mathbf{U}$ of 0 on which $f_t(\xi)$ is continuous *uniformly for* $t \in \mathbb{R}_+$. If also $f_t(\xi) \to 0$ as $t \to \infty$ for all $\xi \in \mathbf{U}$, the equilibrium is asymptotically stable[25] on $\mathbf{U}$. To show stability (alternatively,

[24] Lyapunov's direct method can be used to investigate the stability of more general invariant sets of dynamical systems.

[25] The linear dynamical system $\dot{x} = Ax$ is asymptotically stable at 0 if and only if A is a Hurwitz matrix.

asymptotic stability), one can try to construct a family of hypersurfaces—disjoint, compact, and connected—that all contain 0 and such that all of the trajectories $f_t(x)$ do not leave [are *directed into*] each hypersurface. The easiest, but not the only, way to construct such a family is to use the level sets $\{x|\, V(x) = \delta\}$ of a continuous function $V : \mathbb{R}^n \to \mathbb{R}$ such that $V(0) = 0$ and $V(x) > 0$ for $x \neq x_e$. Such a V is a generalization of the important example $x^\top Qx, Q \gg 0$ and will be called positive definite, written $V \gg 0$. There are similar definitions for positive semidefinite ($V(x) \geq 0, x \neq 0$) and negative semidefinite ($V(x) \leq 0, x \neq 0$). If $(-V) \gg 0$, write $V \ll 0$ (negative definite).

1.7.1 Lyapunov Functions

A function V defined as C^1 and positive definite on some neighborhood $\mathbf{U}$ of the origin is called a gauge function. Nothing is lost by assuming that each level set is a compact hypersurface. Then for $\delta > 0$, one can construct the regions and boundaries

$$\mathrm{V}_\delta := \{x|\, x \neq 0,\ V(x) \leq \delta\},\ \partial\mathrm{V}_\delta = \{x|\, V(x) = \delta\}.$$

A gauge function is said to be a Lyapunov function for the stability of $f_t(0) = 0$, if there exists a $\delta > 0$ such that if $\xi \in \mathrm{V}_\delta$ and $t > 0$,

$$V(f_t(\xi)) \leq V(\xi), \quad \text{that is, } f_t(\partial\mathrm{V}_\delta) \subseteq \mathrm{V}_\delta.$$

One way to prove asymptotic stability uses stronger inequalities:

$$\text{if for } t > 0 \quad V(f_t(\xi)) < V(\xi) \text{ then } f_t(\mathrm{V}_\delta) \subsetneqq \mathrm{V}_\delta.$$

Under our assumption that gauge function V is C^1, $\partial\mathrm{V}_\delta$ has tangents everywhere.[26] Along an integral curve $f_t(x)$,

$$\dot{V}(x) := \left.\frac{dV(f_t(x))}{dt}\right|_{t=0} = \mathbf{f}V(x) \text{ where } \mathbf{f} := \sum_{i=1}^{n} f_i(x)\frac{\partial}{\partial x_i}.$$

Remark 1.4. The quantity $\mathbf{f}V(x) \in \mathbb{R}$ is called the Lie derivative of V with respect to f at x. The first-order partial differential operator $\mathbf{f}$ is called a vector field corresponding to the dynamical system $\dot{x} = f(x)$; see Section B.3.1.[27] △

[26] One can use piecewise smooth V; Hahn [113] describes Lyapunov theory using a one-sided time derivative along trajectories. For an interesting exposition, see Logemann and Ryan [191].

[27] For dynamical system $\dot{x} = f(x)$, the mapping $f : \mathbb{R}^n \to T\mathbb{R}^n \simeq \mathbb{R}^n$ is also called a vector field. Another notation for $\mathbf{f}V(x)$ is $\nabla V \cdot f(x)$.

Definition 1.12. A real function $V : \mathbb{R}^n \to \mathbb{R}$ is called radially unbounded if $V(x) \to \infty$ as $\|x\| \to \infty$. △

Proposition 1.8. *Suppose $V \gg 0$ is C^1 and radially unbounded; if for some $\delta > 0$, on V_δ both $V(x) > 0$ and $\mathbf{f}V(x) < 0$ then*
(i) if $\delta < \epsilon$ then the level sets of V are nested: $\mathrm{V}_\delta \subsetneq \mathrm{V}_\epsilon$;
(ii) for $t > 0$, $V(f_t(\xi)) < V(\xi)$; and
(iii) $V(f_t(\xi)) \to 0$ as $t \to \infty$.
It can then be concluded that V is a Lyapunov function for asymptotic stability of $\mathbf{f}$ at 0, the trajectories $f_t(x)$ will never leave the region V_δ, and the trajectories starting in V_δ will all approach 0 as $t \to \infty$.

A maximal region $\mathbf{U}$ in which all trajectories tend to 0 is called the basin of stability for the origin. If this basin is $\mathbb{R}^n_*$, then $\delta = \infty$ and 0 is said to be globally asymptotically stable — and must be the only equilibrium state; in that case $f_t(\mathrm{V}_\delta) \to \{0\}$, so the basin is connected and simply connected. There are many methods of approximating the boundaries of stability basins [113]; in principal, the boundary can be approximated by choosing a small ball $\mathcal{B}_\epsilon := \{x|\,\|x\| < \epsilon\}$ and finding $f_{-t}(\mathcal{B}_\epsilon)$ for large t. If f has several equilibria, their basins may be very convoluted and intertwined. Up to this point, the state space could be any smooth manifold $\mathcal{M}^n$, but to talk about global stability of *the* equilibrium, the manifold must be homeomorphic to $\mathbb{R}^n$, since all the points of $\mathcal{M}^n$ are pulled toward the equilibrium by the continuous mappings f_T.

If V is radially unbounded [proper] then $\mathbf{f}V(x) \ll 0$ implies global asymptotic stability. Strict negativity of $\mathbf{f}V$ is a lot to ask, but to weaken this condition one needs to be sure that the set where $\mathbf{f}V$ vanishes is not too large. That is the force of the following version of LaSalle's Invariance principle. For a proof see Khalil [156, Ch. 3].

Proposition 1.9. *If gauge function V is radially unbounded, $\mathbf{f}V(x)$ is negative semidefinite, and the only invariant set in $\{x|\,\mathbf{f}V(x) = 0\}$ is $\{0\}$ then $\{0\}$ is globally asymptotically stable for $\dot{x} = f(x)$.*

Example 1.11. For linear time-invariant dynamical systems $\dot{x} = Ax$, we used (1.17) for a given $P \gg 0$ to find a gauge function $V(x) = x^\top Qx$. Let

$$\dot{V} := x^\top(A^\top Q + QA)x; \text{ if } \dot{V} \ll 0$$

then V decreases along trajectories. The level hypersurfaces of V are ellipsoids and V qualifies as a Lyapunov function for global asymptotic stability. In this case, we have in Proposition 1.4 what amounts to a converse stability theorem establishing the existence of a Lyapunov function; see Hahn [113] for such converses. △

Example 1.12. If $V(x) = x^\top x$ and the transition mapping is given by trajectories of the vector field $\mathbf{e} = x^\top \partial/\partial x$ (Euler's differential operator), then $\mathbf{e}V = 2V$.

If the vector field is $\mathbf{a} = x^\top A^\top \partial/\partial x$ with $A^\top = -A$ (a rotation), then $\mathbf{a}V = 0$. That dynamical system then evolves on the sphere $x^\top x = \xi^\top \xi$, to which this vector field is tangent. More generally, one can define dynamical systems on manifolds (see Section B.3.3) such as invariant hypersurfaces. The continuous-time dynamical systems $\dot{x} = f(x)$ that evolve on a hypersurface $V(x) = c$ in $\mathbb{R}^n$ are those that satisfy $\mathbf{f}V = 0$. △

Example 1.13. The matrix $A := \left(\begin{smallmatrix} 0 & 1 \\ -1 & -2 \end{smallmatrix}\right)$ has eigenvalues $\{-1, -1\}$, so 0 is globally asymptotically stable for the dynamical system $\dot{x} = Ax$. A reasonable choice of gauge function is $V(x) = x^\top x$ for which $\dot{V}(x) = -4x_2^2$ which is negative semidefinite, so we cannot immediately conclude that V decreases. However, V is radially unbounded; apply Proposition 1.9: on the set where $\dot{V}$ vanishes, $\dot{x}_2 \neq 0$ except at the origin, which is the only invariant set. △

1.7.2 Lyapunov Exponents*

When the values of the input u of a bilinear system (3.1) satisfy a uniform bound $\|u(t)\| \leq \rho$ on $\mathbb{R}_+$, it is possible to study stability properties by obtaining growth estimates for the trajectories. One such estimate is the Lyapunov exponent (also called Lyapunov number) for a trajectory

$$\lambda(\xi, u) := \limsup_{t\to\infty} \frac{1}{t} \ln(\|X(t;u)\xi\|) \text{ and} \tag{1.49}$$

$$\lambda(u) := \sup_{\|\xi\|=1} \limsup_{t\to\infty} \frac{1}{t} \ln(\|X(t;u)\|), \tag{1.50}$$

the maximal Lyapunov exponent. It is convenient to define, for any square matrix Y, the number $\lambda_{\max}(Y)$ as the maximum of the real part of $\operatorname{spec}(Y)$. Thus if $u = 0$ then $\lambda(\xi, 0) = \lambda_{\max}(A)$, which is negative if A is a Hurwitz matrix. For any small ρ, the Lyapunov spectrum for (3.1) is defined as the set

$$\{\lambda(\xi, u) \,|\, \|u(t)\| \leq \rho, \|\xi\| = 1\}.$$

Example 1.14. Let $\dot{X} = AX + \cos(t)BX,\ X(0) = I$. By Floquet's Theorem (see [117] and [146]), the transition matrix can be factored as $X(t) = F(t)\exp(tG)$ where F has period 2π and G is constant. Here the Lyapunov exponent is $\lambda_{\max}(G)$, whose computation takes some work. In the special case $AB = BA$ with $u(t) = \cos(t)$, it is easy:

$$X(t) = \exp\left(\int_0^t (A + \cos(s)B)\, ds\right) = e^{tA} e^{\sin(t)B};\ \lambda(u) = \lambda_{\max}(A). \qquad \triangle$$

Remark 1.5 (Linear flows).* For some purposes, bilinear systems can be treated as dynamical systems (so-called linear flows) on vector bundles; a few examples must suffice here. If for (1.22) the controls u satisfy some autonomous differential equation $\dot{u} = f(u)$ on Ω, the combined (now autonomous) system becomes a dynamical system on the vector bundle $\Omega \times \mathbb{R}^n \to \Omega$ (see [63, Ch. 6]). Another vector bundle approach was used in [63] and by Wirth [289]: the bundle is $\mathcal{U} \times \mathbb{R}^n \to \mathcal{U}$ where the base (the input space $\mathcal{U}$) is a shift-invariant compact metric space, with transition mapping given by the time shift operator $S\sigma$ acting on $\mathcal{U}$ and (1.22) acting on the fiber $\mathbb{R}^n$. Such a setup is used in Wirth [289] to study the stability of a switching system (1.27) with a compact family of matrices $\mathcal{F}$ and some restrictions on switching times; control-dependent Lyapunov functions are constructed and used to obtain the system's Lyapunov exponents from the Floquet exponents obtained from periodic inputs.

△

1.8 From Continuous to Discrete

Many of the well-known numerical methods for time-variant ordinary differential equations can be used to study bilinear control systems experimentally. These methods provide discrete-time control systems (not necessarily bilinear) that approximate the trajectories of (1.28) and are called *discretizations* of (1.28). Those enumerated in this section are implementable in many software packages and are among the few that yield discrete-time control systems of theoretical interest. Method 2 will provide examples of discrete-time bilinear systems for Chapter 4; this and Method 3 will appear in a stochastic context in Section 8.3. For these examples, let $\dot{x} = Ax + uBx$ with a bounded control u that (to avoid technicalities) is continuous on $[0, T]$; suppose that its actual trajectory from initial state ξ is $X(t; u)\xi,\ t \geq 0$. Choose a large integer N and let $\tau := T/N$. Define a discrete-time input history by $v_N(k) := u(k\tau), 1 \leq k \leq N$. The state values are to be computed only at the times $\{k\tau,\ 1 \leq k \leq N\}$.

1. Sampled-data control systems are used in computer control of industrial processes. A sampled-data control system that approximates $\dot{x} = Ax + uBx$ is in its simplest form

$$x(k\tau; v) = \prod_{j=k}^{1} \exp\Big(\tau\big(A + v(j)B\big)\Big)x. \tag{1.51}$$

   It is exact for the piecewise constant control, and its error for the original continuous control is easily estimated. As a discrete-time control system (1.51) is linear in the state but exponential in v. For more about (1.51), see Preposition 3.15 and 5.3.

2. The simplest numerical method to implement for ordinary differential equations is Euler's discretization

$$x_N(k+1) = (I + \tau(A + v_N(k)B)x_N(k). \tag{1.52}$$

$$\text{For } \epsilon > 0\ \exists u\ \exists N : \max_{k\le N} \|x_N(k) - X(k\tau; u)\xi\| < \epsilon \tag{1.53}$$

is a theorem, essentially the product integral method of Section 1.5.1. Note that (1.52) *is bilinear* with drift term $I + \tau A$. The spectrum of this matrix need not lie in the unit disc even if A is a Hurwitz matrix. Better approximations to (1.28) can be obtained by implicit methods.

3. With the same setup, the midpoint method employs a midpoint state $\bar{x}$ defined by $\bar{x} := \frac{1}{2}(x + \overset{+}{x})$ and is given by an *implicit* relation

$$\overset{+}{x} = x + \frac{\tau}{2}A(v(k))\bar{x};\ \ \text{that is, } \bar{x} = x + \frac{\tau}{2}A(v(k))\bar{x},\ \overset{+}{x} = \bar{x} + \frac{\tau}{2}A(v(k))\bar{x}.$$

$$\text{Explicitly, } x(k+1) = \left(I + \frac{\tau}{2}A(v(k))\right)\left(I - \frac{\tau}{2}A(v(k))\right)^{-1} x(k). \tag{1.54}$$

The matrix inversion is often implemented by an iterative procedure; for efficiency use its value at time k as a starting value for the iteration at $k+1$. The transition matrix in (1.54) is a Padé approximant to the exponential matrix used in (1.51) and again is rational in u. If A is a Hurwitz matrix then the spectrum of $(I + \frac{\tau}{2}A)(I - \frac{\tau}{2}A)^{-1}$ is contained in the unit disc. See Example 1.15 for another virtue of this method. △

Example 1.15. Given that $V(x) := x^\top Qx$ is invariant for a dynamical system $\dot{x} = f(x)$ on $\mathbb{R}^n$, i.e., $\mathbf{f}V(x) = 0 = 2x^\top Qf(x)$, then V is invariant for approximations by the midpoint method (1.54). The proof is that the midpoint discretization obeys $x = \bar{x} - \frac{\tau}{2}f(\bar{x})$, $\overset{+}{x} = \bar{x} + \frac{\tau}{2}f(\bar{x})$. Substituting these values in $V(\overset{+}{x}) - V(x)$, it vanishes for all x and τ. △

Multistep (Adams) methods for a differential equation on $\mathbb{R}^n$ are difference equations of some dimension greater than n. Euler's method (1.52) is the only discretization that provides a bilinear discrete-time system yet preserves the dimension of the state space. Unfortunately Euler's method can give misleading approximations to (1.28) if $|u\tau|$ is large: compare the trajectories of $\dot{x} = -x$ with the trajectories of its Euler approximation $\overset{+}{x} = x - \tau x$. This leads to the peculiar controllability properties of the Euler method discussed in Section 4.5.

Both Euler and midpoint methods will come up again in Section 8.3 as methods of solving stochastic differential equations.

1.9 Exercises

Exercise 1.2. To illustrate Proposition 1.2, show that if

$$A = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix} \text{ then } e^{tA} = e^{ta}\left(1 - at + \frac{a^2t^2}{2}\right)I + e^{ta}(t - at^2)A + e^{ta}\frac{t^2}{2}A^2; \text{ if}$$

$$B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \text{ then } e^{tB} = e^{ta}\frac{b\cos(tb) - a\sin(tb)}{b}I + e^{ta}\frac{\sin(tb)}{b}B. \qquad \triangle$$

Exercise 1.3. The product of exponentials of real matrices is not necessarily the exponential of any real matrix. Show that there is no real solution Y to the equation $e^Y = X$ for

$$X := e^{\begin{bmatrix} 0 & -\pi \\ \pi & 0 \end{bmatrix}} e^{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}} = \begin{bmatrix} -e & 0 \\ 0 & -e^{-1} \end{bmatrix} \qquad \triangle$$

Exercise 1.4. Prove the statements made in Remark 1.1. Reduce the problem to the use of the series (1.6) for a single $k \times k$ Jordan block for the imaginary root $\lambda := \omega\sqrt{-1}$ of multiplicity k which is of the form $\lambda I_k + T$ where the triangular matrix T satisfies $T^k = 0$. △

Exercise 1.5. Even if all n of the eigenvalues of a time-variant matrix $F(t)$ are in the open left half-plane, that is no guarantee that $\dot{x} = F(t)x$ is stable at 0. Consider

$$\dot{x} = (A + u_1B_1 + u_2B_2)x, \text{ where } u_1(t) = \cos^2(t),\ u_2(t) = \sin(t)\cos(t),$$
$$A = \begin{bmatrix} -1 & 1 \\ -1 & \frac{1}{2} \end{bmatrix},\ B_1 = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & -\frac{3}{2} \end{bmatrix},\ B_2 = \begin{bmatrix} 0 & -\frac{3}{2} \\ -\frac{3}{2} & 0 \end{bmatrix}.$$

Its solution for initial value ξ is $X(t;u)\xi$ where

$$X(t;u) = \begin{bmatrix} e^{t/2}\cos(t) & e^{-t}\sin(t) \\ -e^{t/2}\sin(t) & e^{-t}\cos(t) \end{bmatrix}. \tag{1.55}$$

Show that the eigenvalues of the matrix $A + u_1B_1 + u_2B_2$ are constant with negative real parts, for all t, but the solutions $X(t;u)\xi$ grow without bound for almost all ξ. For similar examples and illuminating discussion, see Josić and Rosenbaum [146]. △

Exercise 1.6 (Kronecker products). First read Section A.3.2. Let $\nu = n^2$ and *flatten* the $n \times n$ matrix X (that means, list the $\nu := n^2$ entries of X as a column vector $X^\flat = \mathrm{col}(x_{1,1}, x_{2,1}, \dots, x_{n,n-1}, x_{n,n}) \in \mathbb{R}^\nu$). Using the Kronecker product $\otimes$, show that $\dot{X} = AX$, $X(0) = I$ can be written as the following dynamical system in $\mathbb{R}^\nu$:

$$\dot{X}^\flat = (I \otimes A)X^\flat,\ X^\flat(0) = I^\flat \text{ where } I \otimes A = \mathrm{diag}(A, \dots, A). \qquad \triangle$$

Exercise 1.7 (Kronecker sums). With notation as in Exercise 1.6, take $A \in \mathbb{F}^{n\times n}$ and $\nu := n^2$. The Kronecker sum of A^* and A is (as in Section A.3.3)

$$A^* \boxplus A := I \otimes A^* + A^\top \otimes I.$$

The Lyapunov equation $\mathrm{Ly}_A(Q) = -I$ is now $(A^* \boxplus A)Q^\flat = -I^\flat$; compare (A.11). If the $\nu \times \nu$ matrix $A^* \boxplus A$ is nonsingular, $Q^\flat = -(A^* \boxplus A)^{-1} I^\flat$. Show that if A is upper [lower] triangular with $\mathrm{spec}(A) = \{\alpha_1, \dots, \alpha_n\}$, then $A^* \boxplus A$ is upper [lower] triangular (see Theorem A.2). Show that $A^* \boxplus A$ is invertible if and only if $\alpha_i + \bar{\alpha}_j \neq 0$ for all i, j. When is $\dot{X} = AX$ stable at $X(t) = 0$? △

Exercise 1.8. Discuss, using Section A.3, the stability properties on $\mathbb{R}^{n\times n}$ of $\dot{X} = AX + XB$, $X(0) = Z$ supposing $\mathrm{spec}(A)$ and $\mathrm{spec}(B)$ are known. (Obviously $X(t) = \exp(tA) Z \exp(tB)$.) By Theorem A.2, there exist two unitary matrices S, T such that $A_1 := S^* A S$ is upper triangular and $B_1 := T^* B T$ is lower triangular . Then $\dot{X} = AX + XB$ can be premultiplied by S and postmultiplied by T; letting $Y = SXT$, show that $\dot{Y} = A_1 Y + Y B_1$. The flattened form of this equation (on $\mathbb{C}^\nu$)

$$\dot{Y}^\flat = (A_1 \otimes I + I \otimes B_1^\top) Y^\flat$$

has an upper triangular coefficient matrix. △

Exercise 1.9. Suppose $\dot{x} = Ax$ on $\mathbb{R}^n$ and $z := x \otimes x \in \mathbb{R}^\nu$. Show that $\dot{z} = (A \otimes I_n + I_n \otimes A)z$, using (A.6). △

Exercise 1.10. If $Ax = \alpha x$ and $By = \beta y$, then the $n^2 \times 1$ vector $x \otimes y$ is an eigenvector of $A \otimes B$ with eigenvalue $\alpha\beta$. However, if A and B share eigenvectors then $A \otimes B$ may have eigenvectors that are not such products. If A is a 2×2 matrix in Jordan canonical form, find all the eigenvectors — products and otherwise — of $A \otimes A$. △

Exercise 1.11. Given $\dot{x} = Ax + bu$ with $\{A, b\}$ a controllable pair, let $(z_1, \dots, z_n)$ be the new coordinates with respect to the basis $\{A^{n-1}b, \dots, Ab, b\}$. Show that in these coordinates the system is $\dot{z} = Fz + gu$ with

$$F = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_3 & \cdots & -a_{n-1} \end{bmatrix}, \quad g = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$ △

Exercise 1.12. Derive the Kalman condition (1.43) for controllability from (1.42) with $m = 1$, $B = b$ using the input $u(t) = b^\top \exp(-tA^\top)c$ and $\zeta = 0$. Show that a unique vector c exists for any ξ if and only if (1.43) holds. △

Exercise 1.13. For the discrete-time linear control system on $\mathbb{R}^n$

$$\overset{+}{x} = Ax + ub, \ \ \Omega = \mathbb{R}, x(0) = \xi,$$

find the formula for $x(t)$ and show that the system is controllable if and only if $\operatorname{rank}\left[b\ Ab\ \dots\ A^{n-1}b\right] = n$. △

Exercise 1.14 (Stabilization by feedback). If $\dot{x} = Ax + bu$ is controllable, show using (1.43) and Exercise 1.11 that by an appropriate choice of $c^\top$ in a feedback control $u = c^\top x$ any desired values can be assigned to the n eigenvalues in $\operatorname{spec}(A + bc^\top)$. △

Exercise 1.15. Show that the biaffine control system $\dot{x}_1 = 1 + ux_2$, $\dot{x}_2 = -ux_1$ is controllable on $\mathbb{R}^2$ with unbounded $u \in \mathcal{PK}$. Hint: look at trajectories consisting of a line segment and an arc of a circle. △

Exercise 1.16. There are many simple examples of uncontrollability in the presence of constraints, such as $\dot{x} = -x + u$ with $\Omega = [-1, 1]$; find the set $\mathcal{A}(0)$ attainable from 0. △

Exercise 1.17. Variable structure systems of the form $\dot{x} = F(u)x$ where F is a polynomial in u are good models for some process control problems. Example: Find the attainable set for

$$\dot{x}_1 = ux_1,\ \dot{x}_2 = u^2x_2,\ x(0) = \begin{bmatrix}1\\0\end{bmatrix},\ u(\cdot) \in \mathbb{R}.$$

This can be regarded as a bilinear system with control $(u,\ v) \in \mathbb{R}^2$ where $\Omega := \{(u, v) \mid v = u^2\}$. △

Chapter 2
Symmetric Systems: Lie Theory

2.1 Introduction

Among bilinear control systems, unconstrained symmetric systems have the most complete theory. By default, controls are piecewise constant:

$$\dot{x} = \sum_{i=1}^{m} u_i B_i x, \;\; x \in \mathbb{R}^n_*, \; u \in \mathcal{PK}, \; \Omega = \mathbb{R}^m. \tag{2.1}$$

Symmetric bilinear systems are invariant under time-reversal and scaling: if on $[0, T]$ we replace t with $-t$ and $u(t)$ with $u(T-t)$ then as a control system (2.1) is not changed; and for any $c > 0$ the replacement of t with t/c and u with cu leaves (2.1) unchanged.

Assume always that the matrices in the set $\mathbf{B}^m := \{B_1, \dots, B_m\}$ are linearly independent. The matrix control system in $\mathbb{R}^{n\times n}$ corresponding to (2.1) is

$$\dot{X} = \sum_{i=1}^{m} u_i B_i X, \; u \in \mathcal{PK}, \; X(0) = I. \tag{2.2}$$

The set of $\boldsymbol{\Phi}$ of transition matrices for (2.2) was shown to be a group in Proposition 1.7. It will be shown to be a *Lie group*[1] $\mathbf{G}$ in Section 2.3. First it will be necessary (Section 2.2) to introduce Lie algebras and define the matrix Lie algebra $\mathfrak{g}$ associated with (2.2); their relationship, briefly, is that $\exp(\mathfrak{g})$ generates $\mathbf{G}$.

Matrix Lie groups make possible a geometric treatment of bilinear systems, which will be developed along the following lines. Instead of looking at the individual trajectories of (2.1), one may look holistically at the *action*

[1] Lie groups and Lie algebras are named for their progenitor, Sophus Lie (Norwegian, 1842–1899; pronounced "Lee"). The graduate text of Varadarajan [282] is recommended. An often-cited proof that $\boldsymbol{\Phi}$ is a Lie group was given by Jurdjevic and Sussmann [151, Th. 5.1].

D.L. Elliott, *Bilinear Control Systems*, Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_2, © Springer Science+Business Media B.V. 2009

(Definition B.10) of **G** on $\mathbb{R}^n_*$. The set of states $\mathbf{G}\xi := \{P\xi | P \in \mathbf{G}\}$ is an *orbit* of **G** (Section B.5.5). A set $U \subset \mathbb{R}^n$ is invariant under **G** if $\mathbf{G}U \subset U$, and as a group action the mapping is obviously one-to-one and onto. Therefore, U is the union of orbits. Controllability of system (2.1) on $\mathbb{R}^n_*$ means that **G** acting on $\mathbb{R}^n$ has only two orbits, $\{0\}$ and $\mathbb{R}^n_*$. In the language of Lie groups, **G** is transitive on $\mathbb{R}^n_*$, or with an abuse of language, transitive since $\mathbb{R}^n_*$ is the usual object on which we consider controllability.

A benefit of using group theory and some elementary algebraic geometry in Section 2.5 is that the attainable sets for (2.1) can be found by algebraic computations. This effort also provides necessary algebraic conditions for the controllability of bilinear systems with drift, discussed in the next chapter.

Contents of this Chapter

Please refer to the basic definitions and theorems concerning Lie algebras given in Appendix B. Section 2.2 introduces matrix Lie algebras and the way the set $\mathbf{B}^m$ of control matrices generates $\mathfrak{g}$. In particular, Section 2.2.3 contains Theorem 2.1 (Campbell–Baker–Hausdorff). Section 2.2.4 provides an algorithm `LieTree` that produces a basis for $\mathfrak{g}$.

After the introduction to matrix Lie groups[2] in Section 2.3, in Section 2.4 their orbits on $\mathbb{R}^n$ are studied. The *Lie algebra rank condition* defined in Section 2.4.2 is necessary and sufficient for transitivity; examples are explored. Section 2.5 introduces computational algebraic geometry methods (see Sections C.2 and C.2.1) that lead in Section 2.5.2 to a payoff — new algebraic tests that can be used to find orbits of **G** on $\mathbb{R}^n_*$ and decide whether or not **G** is transitive.

The system (2.2) can be used to study the action of **G** on its homogeneous spaces; these are defined and discussed in Section 2.7. Sections 2.8 and 2.9 give examples of canonical coordinates for Lie groups and their use as formulas for the representation and computation of transition matrices of (2.2). Sections 2.10 and 2.11 are more specialized — respectively in complex bilinear systems (with a new significance for weak transitivity in 2.10.2) and how to generate any of the transitive Lie algebras with two matrices. Section 2.12 contains exercises.

2.2 Lie Algebras

Matrix multiplication in $\mathbb{F}^{n\times n}$ is not commutative but obeys the associative law $A(BC) = (AB)C$. Another important binary operation, the Lie bracket $[A, B] := AB - BA$, is not associative. It comes up naturally in control theory

[2] For a careful development of matrix Lie groups (not necessarily closed in $GL(n, \mathbb{R})$), see Rossman [225] (2005 printing or later) where they are called *linear groups.*

because the result of applying two controls A and B in succession depends on the order in which they are applied.

Example 2.1. Let the matrix control system $\dot{X} = (u_1 A + u_2 B)X$ have piecewise constant input $u(t)$, $0 \le t \le 4\tau$ with components

$$\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} := \frac{-1}{\sqrt{\tau}}\begin{bmatrix} 0 \\ 1 \end{bmatrix},\ 0 \le t < \tau;\ \frac{-1}{\sqrt{\tau}}\begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \tau \le t < 2\tau;$$
$$\frac{1}{\sqrt{\tau}}\begin{bmatrix} 0 \\ 1 \end{bmatrix},\ 2\tau \le t < 3\tau;\ \frac{1}{\sqrt{\tau}}\begin{bmatrix} 0 \\ 1 \end{bmatrix},\ 3\tau \le t \le 4\tau.$$

Using this input u, we obtain the transition matrix

$$X(4\tau; u) = \mathrm{e}^{\sqrt{\tau}A}\mathrm{e}^{\sqrt{\tau}B}\mathrm{e}^{-\sqrt{\tau}A}\mathrm{e}^{-\sqrt{\tau}B}.$$

Using the first few terms of the Taylor series for the matrix exponential in each of these four factors,

$$X(4\tau; u) = I + \tau[A, B] + \frac{\tau^{3/2}}{2}([A,[A,B]] + [B,[A,B]]) + \tau^2 R; \tag{2.3}$$
$$[A, B] = \lim_{\tau \downarrow 0} \frac{1}{\tau}(\mathrm{e}^{\sqrt{\tau}A}\mathrm{e}^{\sqrt{\tau}B}\mathrm{e}^{-\sqrt{\tau}A}\mathrm{e}^{-\sqrt{\tau}B} - I) = \dot{X}(0; u).$$

For more about the remainder R in (2.3) see Exercise (2.1). △

Equipped with the Lie bracket $[A, B] := AB - BA$ as its multiplication, the linear space $\mathbb{F}^{n\times n}$ is a nonassociative algebra denoted by $\mathfrak{gl}(n, \mathbb{F})$. It is easily checked that it satisfies Jacobi's axioms: for all A, B, C in $\mathfrak{gl}(n, \mathbb{F})$ and α, β in $\mathbb{F}$

linearity:	$[A, \alpha B + \beta C] = \alpha[A, B] + \beta[A, C],\ \alpha, \beta \in \mathbb{F};$	(Jac.1)
antisymmetry:	$[A, B] + [B, A] = 0;$	(Jac.2)
Jacobi identity:	$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0.$	(Jac.3)

Therefore $\mathfrak{gl}(n, \mathbb{F})$ is a Lie algebra as defined in Section B.1; please see that section at this point. Lie algebras are denoted by letters $\mathfrak{c}, \mathfrak{d}, \mathfrak{e}, \mathfrak{f}, \mathfrak{g}$, etc. from the *fraktur* font, often chosen to reflect their definitions. Thus $\mathfrak{gl}(n, \mathbb{F})$ is a general linear Lie algebra. Lie algebra theory is a mathematically deep subject — see Jacobson [143] or Varadarajan [282] for evidence toward that statement — but the parts needed for an introduction to bilinear systems are easy to follow.

Definition 2.1. A matrix Lie algebra is a Lie subalgebra $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{F})$. △

Example 2.2. The standard basis for $\mathfrak{gl}(n, \mathbb{F})$ is the set of elementary matrices: an elementary matrix $1^{(i,j)}$ has 1 in the (i, j) position and 0 elsewhere.

Since $1^{(i,j)}1^{(k,l)} = \delta_{(j,k)}1^{(i,l)}$, the Lie bracket multiplication table for $\mathfrak{gl}(n,\mathbb{F})$ is $[1^{(i,j)}, 1^{(k,l)}] = \delta_{j,k}1^{(i,l)} - \delta_{l,i}1^{(k,j)}$. △

2.2.1 *Conjugacy and Isomorphism*

Given a nonsingular matrix $P \in \mathbb{F}^{n\times n}$, a linear coordinate transformation is a one-to-one linear mapping $x = Py$ on $\mathbb{R}^n$; it takes $\dot{x} = Ax$ into $\dot{y} = P^{-1}APy$. It satisfies

$$[P^{-1}AP, P^{-1}BP] = P^{-1}[A, B]P,$$

so it takes a matrix Lie algebra $\mathfrak{g}$ to an isomorphic[3] matrix Lie algebra $\mathfrak{h}$ related by $P\mathfrak{g} = \mathfrak{h}P$; $\mathfrak{g}$, $\mathfrak{h}$ are called conjugate Lie algebras. Given real $\mathfrak{g}$, there are often reasons (the use of canonical forms) to employ a complex coordinate transformation P. The resulting space of complex matrices $\mathfrak{h}$ is still a real Lie algebra; it is closed under $\mathbb{R}$-linear operations but not $\mathbb{C}$-linear.

Another example of a Lie algebra isomorphism relates a matrix Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(n,\mathbb{R})$ with a Lie algebra of linear vector fields $\boldsymbol{\mathfrak{g}} \subset \mathcal{V}(\mathbb{R}^n)$ (Section B.3.1). With $A, B \in \mathbb{R}^{n\times n}$, let

$$\mathbf{a} := (Ax)^\top \frac{\partial}{\partial x}, \ \mathbf{b} := (Bx)^\top \frac{\partial}{\partial x}, \text{ then}$$

$$A \simeq \mathbf{a},\ B \simeq \mathbf{b},\ [A, B] \simeq \mathbf{ba} - \mathbf{ab} = [\mathbf{a}, \mathbf{b}] \tag{2.4}$$

using the Lie bracket for vector fields defined in Section B.3.5. The matrix Lie algebra $\mathfrak{g}$ and vector field Lie algebra $\boldsymbol{\mathfrak{g}}$ are isomorphic; in particular, $\{A, B\}_{\mathcal{L}} \simeq \{\mathbf{a}, \mathbf{b}\}_{\mathcal{L}}$.

2.2.2 *Some Useful Lie Subalgebras*

If a set of matrices $\mathfrak{L}$ has the property that the Lie bracket of any two of its elements belongs to $\mathrm{span}_{\mathbb{R}}(\mathfrak{L})$, then $\mathfrak{L}$ will be called real-involutive.[4] If real-involutive $\mathfrak{L}$ is a linear space it is, of course, a Lie algebra.

The linear space of all those real matrices whose trace is zero is real-involutive, so it is a matrix Lie algebra: $\mathfrak{sl}(n,\mathbb{R})$, the special linear algebra. For this example and more see Table 2.1. It gives six linear properties that a matrix may have, each preserved under Lie bracketing and linear combination, and the maximal Lie subalgebra of $\mathfrak{gl}(n,\mathbb{R})$ with each property.

---

[3] See Section B.1.3.

[4] A caveat about *involutive*: a family $\mathcal{F}$ of smooth vector fields is called involutive if given any $\mathbf{f}, \mathbf{g} \in \mathcal{F}$, $[\mathbf{f}, \mathbf{g}]$ belongs to the span of $\mathcal{F}$ over the ring of smooth functions.

$(a)\ x_{i,j} = 0,\ i \neq j,$	$\mathfrak{d}(n, \mathbb{R}),\ \ell = n;$
$(b)\ X^\top + X = 0,$	$\mathfrak{so}(n),\ \ell = n(n-1)/2;$
$(c)\ x_{i,j} = 0,\ i \leq j,$	$\mathfrak{n}(n, \mathbb{R}),\ \ell = n(n-1)/2;$
$(d)\ x_{i,j} = 0,\ i < j,$	$\mathfrak{t}^u(n, \mathbb{R}),\ \ell = n(n+1)/2;$
$(e)\ \operatorname{tr}(X) = 0,$	$\mathfrak{sl}(n, \mathbb{R}),\ \ell = n^2 - 1;$
$(f)\ XH = 0,$	$\mathfrak{fix}(H),\ \ell = n^2 - \operatorname{rank} H.$

Table 2.1. Some matrix Lie algebras and their dimensions. For (f), $H \in \mathbb{R}^{n\times n}$ has rank less than n. Cf. Table 2.3.

Lie algebras with Property (c) are nilpotent Lie algebras (Section B.2). Those with Property (f) are (obviously) Abelian. Property (d) implies solvability; see Section B.2.1 for basic facts. Section 2.8.2 discusses transition matrix representations for solvable and nilpotent Lie algebras.

Any similarity $X = P^{-1}YP$ provides an isomorphism of the orthogonal algebra $\mathfrak{so}(n)$ to $\{Y|\ Y^\top Q + QY = 0\}$ where $Q = PP^\top \gg 0$.

However, suppose $Q = Q^\top$ is nonsingular but is not sign definite with sign pattern (see page 7) $\{\mathfrak{p}_Q, \mathfrak{z}_Q, \mathfrak{n}_Q\} = \{p, 0, q\}$; by its signature is meant the pair (p, q). The resulting Lie algebras are

$$\mathfrak{so}(p, q) := \left\{X|\ X^\top Q + QX = 0,\ Q^\top = Q\right\};\ \mathfrak{so}(p, q) = \mathfrak{so}(q, p). \tag{2.5}$$

An example (used in relativistic physics) is the Lorentz algebra $\mathfrak{so}(3, 1) \subset \mathfrak{gl}(4, \mathbb{R})$, for which $Q = \operatorname{diag}(-1, 1, 1, 1)$.

Remark 2.1. Any real abstract Lie algebra has at least one faithful representation as a matrix Lie algebra on some $\mathbb{R}^n$. (See Section B.1.5 and Ado's Theorem B.2.) Sometimes different (nonconjugate) faithful representations can be found on the same $\mathbb{R}^n$; for some examples of this, see Section 2.6.1. △

Exercise 2.1. The remainder R in (2.3) need not belong to $\{A, B\}_{\mathcal{L}}$. For a counterexample, use the elementary matrices $A = 1^{(1,2)}$ and $B = 1^{(2,1)}$, which generate $\mathfrak{sl}(2, \mathbb{R})$, in (2.3); show $\operatorname{tr} R = t^4$. △

2.2.3 *The Adjoint Operator*

For any matrix $A \in \mathfrak{g}$, the operation $\operatorname{ad}_A(X) := [A, X]$ on $X \in \mathfrak{g}$ is called the adjoint action of A. Its powers are defined by

$$\operatorname{ad}_A^0(X) = X,\ \operatorname{ad}_A^1(X) = [A, X];\ \operatorname{ad}_A^k(X) = [A, \operatorname{ad}_A^{k-1}(X)],\ k > 0.$$

Note the linearity in both arguments: for all $\alpha, \beta \in \mathbb{F}$

$$\operatorname{ad}_A(\alpha X + \beta Y) = \alpha \operatorname{ad}_A(X) + \beta \operatorname{ad}_A(Y) = \operatorname{ad}_{\alpha A + \beta B}(X). \tag{2.6}$$

Here are some applications of ad_X:

(i) the set of matrices that commute with matrix B is a Lie algebra,

$$\mathfrak{ker}(\mathrm{ad}_B) := \{X \in \mathfrak{gl}(n, \mathbb{R}) |\ \mathrm{ad}_B(X) = 0\};$$

(ii) for the *adjoint representation* $\mathrm{ad}_{\mathfrak{g}}$ of a Lie algebra $\mathfrak{g}$ on itself see Section B.1.8;

(iii) the Lie algebra $\mathfrak{c} := \{X \in \mathfrak{gl}(n, \mathbb{R}) |\ \mathrm{ad}_{\mathfrak{g}}(X) = 0\}$, the kernel of $\mathrm{ad}_{\mathfrak{g}}$ in $\mathfrak{g}$, is called the center of $\mathfrak{g}$.

In addition to its uses in Lie algebra, ad_A has uses in matrix analysis. Given any $A \in \mathbb{F}^{n\times n}$, the mapping $\mathrm{ad}_A(X) = [A, X]$ is defined for all $X \in \mathbb{F}^{n\times n}$. If as in Sections A.3 and 1.9 we map a matrix $X \in \mathbb{F}^{n\times n}$ to a vector $X^\flat \in \mathbb{F}^\nu$, where $\nu := n^2$, the $\flat$ mapping induces a representation of the operator ad_A on $\mathbb{F}^\nu$ by a $\nu \times \nu$ matrix (see Section A.3.3)

$$\mathrm{ad}^\flat_A := I_n \otimes A - A^\top \otimes I_n; \quad \mathrm{ad}^\flat_A X^\flat = [A, X]^\flat.$$

The dimension of the nullspace of ad_A on $\mathbb{F}^{n\times n}$ is called the nullity of ad_A and is written $\mathrm{null}(\mathrm{ad}_A)$; it equals the dimension of the nullspace of $\mathrm{ad}^\flat_A$ acting on $\mathbb{F}^\nu$.

Proposition 2.1. *Given that* $\mathrm{spec}(A) = \{\alpha_1, \dots, \alpha_n\}$,

(i) $\mathrm{spec}(\mathrm{ad}_A) = \{\alpha_i - \alpha_j :\ i, j \in 1\,..\,n\}$ and
(ii) $n \le \mathrm{null}(\mathrm{ad}_A) \le \nu$.

Proof. Over $\mathbb{C}$, to each eigenvalue α_i corresponds a unit vector x_i such that $Ax_i = \alpha_i x_i$. Then $\mathrm{ad}_A\ x_i x_j^\top = (\alpha_i - \alpha_j) x_i x_j^\top$. If A has repeated eigenvalues and fewer than n eigenvectors, $\mathrm{ad}_A(X) = \lambda X$ will have further matrix solutions which are not products of eigenvectors of A.

For the details of $\mathrm{null}(\mathrm{ad}_A)$, see the discussion in Gantmacher [101, Ch. VIII]; the structure of the $\mathrm{null}(\mathrm{ad}_A)$ linearly independent matrices that commute with A are given, depending on the Jordan structure of A; $A = I$ achieves the upper bound $\mathrm{null}(\mathrm{ad}_A) = \nu$. The lower bound n is achieved by matrices with n distinct eigenvalues, and in that case the commuting matrices are all powers of A. □

The adjoint action of a Lie group $\mathbf{G}$ on its own Lie algebra $\mathfrak{g}$ is defined as

$$\mathrm{Ad} : \mathbf{G} \times \mathfrak{g} \to \mathfrak{g};\ \mathrm{Ad}_P(X) := PXP^{-1},\ P \in \mathbf{G},\ X \in \mathfrak{g}, \tag{2.7}$$

and $P \mapsto \mathrm{Ad}_P$ is called the adjoint representation of $\mathbf{G}$, see Section B.1.8. The following proposition shows how ad and Ad are related.

Proposition 2.2 (Pullback). *If* $A, B \in \mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{F})$ *and* $P := \exp(tA)$ *then*

$$\mathrm{Ad}_P(B) := \mathrm{e}^{tA} B \mathrm{e}^{-tA} = \mathrm{e}^{t\,\mathrm{ad}_A}(B) \in \mathfrak{g}, \ \ \textit{and}\ \mathrm{e}^{t\,\mathrm{ad}_A}(\mathfrak{g}) = \mathfrak{g}. \tag{2.8}$$

Proof. On $\mathbb{F}^{n\times n}$, the linear differential equation

$$\dot{X} = AX - XA,\ X(0) = B,\ t \geq 0$$

is known to have a unique solution of the form

$$X(t) = e^{tA}Be^{-tA};\ \text{as a formal power series}$$
$$X(t) = e^{t\,\mathrm{ad}_A}(B) := B + t[A,B] + \frac{t^2}{2}[A,[A,B]] + \cdots \tag{2.9}$$

in which every term belongs to the Lie algebra $\{A,B\}_{\mathcal{L}} \subset \mathfrak{g}$. The series (2.9) is majorized by the Taylor series for $\exp(2t\|A\|)\,\|B\|$; therefore it is convergent in $\mathbb{F}^{n\times n}$, hence in $\mathfrak{g}$. Since all this is true for all $B \in \mathfrak{g}$, $\exp(t\,\mathrm{ad}_A)\mathfrak{g} = \mathfrak{g}$.[5] □

Example 2.3. The following way of using the pullback formula[6] is typical. For $i \in 1,\ldots,\ell$, let

$$Q_i := \exp(B_i);\quad \mathrm{Ad}_{Q_i}(Y) := \exp(B_i)Y\exp(-B_i).\ \text{Let}$$
$$F(v) = e^{v_1B_1}e^{v_2B_2}\cdots e^{v_\ell B_\ell};\ \text{ then at } v = 0$$
$$\frac{\partial F}{\partial v_1} = B_1F,\ \frac{\partial F}{\partial v_2} = \mathrm{Ad}_{Q_1}(B_2)F,\ \frac{\partial F}{\partial v_3} = \mathrm{Ad}_{Q_1}(\mathrm{Ad}_{Q_2}(B_3))F,\ldots$$
$$\frac{\partial F}{\partial v_\ell} = \mathrm{Ad}_{Q_1}(\mathrm{Ad}_{Q_2}(\mathrm{Ad}_{Q_3}\cdots(\mathrm{Ad}_{Q_{\ell-1}}(B_\ell))\cdots))F.$$ △

Theorem 2.1 (Campbell–Baker–Hausdorff). *If $\|X\|$ and $\|Y\|$ are sufficiently small then $e^Xe^Y = e^{\mu(X,Y)}$ where the matrix $\mu(X,Y)$ is an element of the Lie algebra $\mathfrak{g} = \{X,Y\}_{\mathcal{L}}$ given by the power series*

$$\mu(X,Y) := X + Y + \frac{[X,Y]}{2} + \frac{[[X,Y],Y] + [[Y,X],X]}{12} + \ldots, \tag{2.10}$$

which converges if $\|X\| + \|Y\| < \dfrac{\log(2)}{\eta}$, *where*

$$\eta = \inf_{A,B\in\mathfrak{g}} \{\delta > 0 \,|\, \|\,[A,B]\,\| \leq \delta\|A\|\cdot\|B\|\}. \tag{2.11}$$

There exists a ball $\mathfrak{B}_K$ of radius $K > 0$ such that if $X, Y \in N_K$ then

$$\|\mu(X,Y) - X - Y\| \leq \eta\|X\|\|Y\| \ \text{ and } \ \|\mu(X,Y)\| \leq 2K + \eta K^2. \tag{2.12}$$

[5] By the Cayley–Hamilton Theorem, $\exp(t\,\mathrm{ad}_A)$ is a polynomial in ad_A of degree at most n^2; for more details see Altafini [6].

[6] The pullback formula (2.8) is often, mistakenly, called the Campbell–Baker–Hausdorff formula (2.10).

2.2.3.1 Discussion*

The Campbell–Baker–Hausdorff Theorem, which expresses $\log(XY)$ as a Lie series in $\log(X)$ and $\log(Y)$ when both X and Y are near I, is proved in Jacobson [143, V.5], Hochschild [129], and Varadarajan [282, Sec. 2.15]. A typical proof of (2.10) starts with the existence of $\log(e^X e^Y)$ as a formal power series in noncommuting variables, establishes that the terms are Lie monomials (Sections B.1.6, 2.2.4), and obtains the estimates necessary to establish convergence for a large class of Lie algebras (Dynkin algebras) that includes all finite-dimensional Lie algebras. For (2.11), see [282, Problem 45] or [129, Sec. X.3]; η can be found from the structure tensor $\gamma^i_{j,k}$ (Section B.1.7) and there are bases for $\mathfrak{g}$ for which $\eta = 1$. The bound on μ in (2.12) is quoted from Hilgert et al. [127, A.1.4].

If the CBH function μ defined on $\mathfrak{g}$ by (2.10), $e^X e^Y = e^{\mu(X,Y)}$, is restricted to neighborhood U of 0 in $\mathfrak{g}$ (U sufficiently small, symmetric, and star-shaped), μ locally defines a product $X * Y = \mu(X, Y) \in \mathfrak{g}$. With respect to U, the multiplication $*$, with $\iota = 0$, is real-analytic and defines a *local Lie group* of matrices. This construction is in Duistermaat and Kolk [77] and is used in Hilgert et al. [127] to build a *local Lie semigroup* theory.

2.2.4 *Generating a Lie Algebra*

To each real m-input, bilinear system corresponds a generated matrix Lie algebra. The matrices in the list $\mathbf{B}^m$, assumed to be linearly independent, will be called the generators of this Lie algebra $\mathfrak{g} := \mathbf{B}^m_{\mathcal{L}}$, the smallest real-involutive linear subspace of $\mathfrak{gl}(n, \mathbb{R})$ containing $\mathbf{B}^m$. It has dimension $\ell \le \nu$. The purpose of this section is to motivate and provide an effective algorithm (`LieTree`) that will generate a basis $\mathbf{B}^\ell$ for $\mathfrak{g}$. It is best for present purposes to include the generators in this list; $\mathbf{B}^\ell = \{B_1, \ldots, B_m, \ldots, B_\ell\}$.

If $P^{-1}\mathbf{B}^m P = \widetilde{\mathbf{B}}^m$ then $\mathbf{B}^m_{\mathcal{L}}$ and $\widetilde{\mathbf{B}}^m_{\mathcal{L}}$ are conjugate Lie algebras; also, two conjugate Lie algebras can be given conjugate bases. The important qualitative properties of bilinear systems are independent of linear coordinate changes but depend heavily on which brackets are of lowest degree. Obtaining control directions corresponding to brackets of high order is costly, as (2.3) suggests.

Specializing to the case $m = 2$ (the case $m > 2$ requires few changes) and paraphrasing some material from Section B.1.6 will help to understand `LieTree`. Define $\mathfrak{L}_2$ as the Lie ring over $\mathbb{Z}$ generated by two indeterminates α, β equipped with the Lie bracket; if $p, q \in \mathfrak{L}_2$ then $[p, q] := pq - qp \in \mathfrak{L}_2$. $\mathfrak{L}_2$ is called the free Lie algebra on two generators. A Lie monomial p is a formal expression obtained from α, β by repeated bracketing *only*; the degree $\deg(p)$ is the number of factors in p, for example

$$\alpha,\ \beta,\ [\alpha,\beta],\ [\alpha,[\alpha,\beta]],\ [[\alpha,\beta],[\alpha,[\beta,[\alpha,\beta]]]],$$

with respective **degrees** 1, 1, 2, 3, 6, are Lie monomials.

A Lie monomial p is called **left normalized** if either (i) $p \in \{\alpha,\beta\}$ or (ii) $p = [\gamma, q]$, where $\gamma \in \{\alpha,\beta\}$ and Lie monomial q, is left normalized.[7] Thus if $\deg(p) = N$, there exist $\gamma \in \{\alpha,\beta\}$ and nonnegative indices j_i such that

$$p = \mathrm{ad}_\alpha{}^{j_1}\mathrm{ad}_\beta{}^{j_2}\dots\mathrm{ad}_\beta{}^{j_k}(\gamma),\ \sum_{i=1}^{k} j_i = N.$$

It is convenient to organize the left normalized Lie monomials into a list of lists $\mathfrak{T} = \{\mathfrak{T}_1, \mathfrak{T}_2, \dots\}$ in which $\mathfrak{T}_d$ contains monomials of degree d. It is illuminating to define $\mathfrak{T}$ inductively as a dyadic tree: level $\mathfrak{T}_d$ is obtained from $\mathfrak{T}_{d-1}$ by ad_α and ad_β actions (in general, by m such actions). The monomials made redundant by Jac.1 and Jac.2 are omitted:

$$\mathfrak{T}_1 := \{\alpha,\ \beta\};\ \mathfrak{T}_2 := \{[\beta,\alpha]\};\ \mathfrak{T}_3 = \{[\alpha,[\beta,\alpha]],[\beta,[\beta,\alpha]]\}; \tag{2.13}$$
$$\mathfrak{T}_i = \{\gamma_{1,i},\gamma_{2,i},\dots\} \Rightarrow\ \mathfrak{T}_{i+1} := \{[\alpha,\gamma_{1,i}],[\beta,\gamma_{1,i}],[\alpha,\gamma_{2,i}],[\beta\gamma_{2,i}],\dots\}.$$

Now concatenate the $\mathfrak{T}_i$ into a simple list

$$\mathfrak{T}_\star := \mathfrak{T}_1 \star \mathfrak{T}_2 \star \cdots = \{\alpha,\beta,[\beta,\alpha],[\alpha,[\beta,\alpha]],[\beta,[\beta,\alpha]],\dots\}.$$

As a consequence of Jac.3, there are even more dependencies, for instance $[\alpha,[\beta,[\alpha,\beta]]] + [\beta,[\alpha,[\alpha,\beta]]] = 0$. The Lie bracket of two elements of $\mathfrak{T}_\star$ is not necessarily left normalized, so it may not be in $\mathfrak{T}_\star$. The module over $\mathbb{Z}$ generated by the left normalized Lie monomials is

$$\mathbb{Z}(\mathfrak{T}_\star) := \Big\{\sum_{j=1}^{k} i_j p_j \Big| i_j \in \mathbb{Z}, k \in \mathbb{N}, p_j \in \mathfrak{T}_\star\Big\}.$$

Proposition 2.3. *If $p, q \in \mathbb{Z}(\mathfrak{T}_\star)$ then $[p,q] \in \mathbb{Z}(\mathfrak{T}_\star)$. Furthermore, $\mathbb{Z}(\mathfrak{T}_\star) = \mathfrak{L}_2$.*

Proof. The first statement follows easily from the special case $p \in \mathfrak{T}_\star$. The proof of that case is by induction on $\deg(p)$, for arbitrary $q \in \mathbb{Z}(\mathfrak{T}_\star)$.

I If $\deg(p) = 1$ then $p = \gamma \in \{\alpha,\beta\}$; so $[\gamma,q] \in [\gamma,\mathbb{Z}(\mathfrak{T}_\star)] = \mathbb{Z}(\mathfrak{T}_\star)$.
II Induction hypothesis: if $\deg(p) = k$ then $[p_k, q] \in \mathbb{Z}(\mathfrak{T}_\star)$.

Induce: let $\deg(p) = k+1$, then $p \in \mathfrak{T}_{k+1}$. By (2.13) (the definition of $\mathfrak{T}_{k+1}$), there exists some $\gamma \in \{\alpha,\beta\}$ and some monomial $c_k \in \mathfrak{T}_k$ such that $p = [\gamma, c_k]$. By Jac.3 $[q,p] = [q,[\gamma,c_k]] = -[\gamma,[c_k,q]] - [c_k,[\gamma,q]]$. Here $[c_k,q] \in \mathbb{Z}(\mathfrak{T}_\star)$ by (II),

[7] Left normalized Lie monomials appear in the linearization theorem of Krener [164, Th. 1].

so $[\gamma, [c_k, q]] \in \mathbb{Z}(\mathfrak{T}_\star)$ by (I). Since $[\gamma, q] \in \mathbb{Z}(\mathfrak{T}_\star)$, by (II) $[c_k, [\gamma, q]] \in \mathbb{Z}(\mathfrak{T}_\star)$. Therefore $[p, q] \in \mathbb{Z}(\mathfrak{T}_\star)$.

From the definition of $\mathbb{Z}(\mathfrak{T}_\star)$, it contains the two generators, each of its elements belongs to $\mathfrak{L}_2$ and it is closed under $\mathbb{Z}$ actions. It is, we have just shown, closed under the Lie bracket. Since $\mathfrak{L}_2$ is the smallest such object, each of its elements belongs to $\mathbb{Z}(\mathfrak{T}_\star)$, showing equality. □

We are given two matrices $A, B \in \mathbb{R}^{n\times n}$; they generate a real Lie algebra $\{A, B\}_{\mathcal{L}}$. There exists a well-defined mapping $h : \mathfrak{L}_2 \to \mathbb{R}^{n\times n}$ such that $h(\alpha) = A$, $h(\beta) = B$, and Lie brackets are preserved (h is a Lie ring homomorphism). The image of h in $\mathbb{R}^{n\times n}$ is $h(\mathfrak{L}_2) \subset \{A, B\}_{\mathcal{L}}$. We can define the tree of matrices $h(\mathfrak{T}) := \{h(\mathfrak{T}_1), h(\mathfrak{T}_2), \dots\}$ and the corresponding list of matrices $h(\mathfrak{T}_\star) = \{A, B, [B, A], \dots\}$. The serial order of the matrices in $h(\mathfrak{T}_\star) = \{X_1, X_2, \dots\}$ lifts to a total order on $h(\mathfrak{T})$.

Call a matrix X_k in $h(\mathfrak{T})$ past-dependent if it is linearly dependent on $\{X_1, \dots, X_{k-1}\}$. In the tree $h(\mathfrak{T})$, the descendants of past-dependent matrices are past-dependent. Construct from $h(\mathfrak{T})$ a second tree of matrices $h(\widehat{T}) = \{h(\widehat{T}_1), \dots\}$ that preserves degree (as an image of a monomial) but has been pruned by removing all the past-dependent matrices.

Proposition 2.4. *$h(T)$ can have no more than ν levels. As a set, $h(\widehat{T})$ is linearly independent and is a basis for $\{A, B\}_{\mathcal{L}} = \operatorname{span}_{\mathbb{R}} h(\widehat{T})$.*

Proof. The real span of $h(\widehat{T})$ is $\operatorname{span}_{\mathbb{R}} h(\mathbb{Z}(\mathfrak{T}_\star)) = \operatorname{span}_{\mathbb{R}} h(\mathfrak{L}_2)$; by Proposition 2.3, this space is real-involutive. Recall that $\{A, B\}_{\mathcal{L}}$ is the smallest real-involutive $\mathbb{R}$-linear space that contains A and B. □

Given $\mathbf{B}^m$, algorithms that will compute a basis for $\mathbf{B}^m_{\mathcal{L}}$ have been suggested in Boothby and Wilson [32] and Isidori [141, Lemma 2.4.1]; here Proposition 2.4 provides a proof for the `LieTree` algorithm implemented in Table 2.2. That algorithm can now be explained.

`LieTree` accepts m generators as input and outputs the list of matrices `Basis`, which is also available as a tree `T`. These matrices correspond to left normalized Lie monomials from the m-adic tree $\mathfrak{T}$.[8] Initially `tree[1]`, ..., `tree[`n^2`]` and `Basis` themselves are empty.

If the input list `gen` of `LieTree` is linearly dependent, the algorithm stops with an error statement; otherwise, `gen` is appended to `Basis` and also assigned to `tree[1]`. At level i the innermost (j, k) loop computes `com`, the Lie bracket of the jth generator with the kth matrix in level i; `trial` is the concatenation of `Basis` and `com`. If `com` is past-dependent (`Dim[trial]=Dim[Basis]`), it is rejected; otherwise, `com` is appended to `Basis`. If level i is empty, the algorithm terminates. That it produces a basis for $\mathfrak{g}$ follows from the general case of Proposition 2.4.

---

[8] Unlike general Lie algebras (nonlinear vector fields, etc.), the bases for matrix Lie algebras are smaller than Philip Hall bases.

```
Rank::Usage="Rank[M]; matrix M(x); output is generic rank."
Rank[M_]:= Length[M[[1]]]-Dimensions[NullSpace[M]][[1]];
Dim::Usage="Dim[L]; L is a list of n by n matrices. \n
Result: dimension of span(L). "
Dim[L_]:=Rank[Table[Flatten[L[[i]]],{i,Length[L]}]];

LieTree::Usage="LieTree[gen,n]; \n
gen=list of n x n matrices; output=basis of \n
the Lie algebra generated by gen."
LieTree[gen_, n_]:= Module[{db,G, i,j,k, newLevel,newdb,lt,
L={},nu=n^2,tree,T,m=Length[gen],r=Dim[gen]},
Basis=gen; If[r<m,Return["Dependent list"]];
LieWord=""; T=Array[tree, nu]; tree[1]=gen;
For[i=2, i<nu-1, i++,{L=tree[i-1];Lm=Length[L];
db=Dim[Basis]; newLevel={};
For[j=1,j<m+1,j++,{For[k=1,k<Lm+1,k++,
{G=gen[[k]];com=G.L[[j]]-L[[j]].G;
trial= Append[Basis, com]; newdb= Dim[trial];
If[newdb >db, {newLevel=Append[newLevel,com];
Basis=Append[Basis, com];
LieWord=StringJoin[LieWord,ToString[{i,j,k}]];
db= newdb;};];};];};];
tree[i]=newLevel; lt=Length[newLevel];
If[lt <1, Break[], r =lt]};]; Basis];
```

Table 2.2. Script `TreeComp`. `LieTree` produces a `Basis` for $\mathbf{B}^m_{\mathcal{L}}$ from gen:= $\mathbf{B}^m$.

The function `Dim` finds the dimensions of (the span of) a list of matrices $\{B_1, \dots, B_l\}$ by flattening the matrices to vectors of dimension ν (see Section A.3) and finding the rank of the $\nu \times l$ matrix $\left[B_1^\flat \cdots B_l^\flat\right]$.

The global variable `LieWord` encodes in a recursive way the left normalized Lie monomials represented in the output list. The triple $\{i, j, k\}$ indicates that an accepted item in level T_i of `Basis` is the bracket of the kth generator with the jth element of level T_{i-1}. Thus if $T_1 = \{A, B\}$, the `LieWord` `{2,1,2}{3,1,1}{3,1,2}` means $T_2 = \{[B, A]\}$ and $T_3 = \{[A, [B, A]], [B, [B, A]]\}$. This encoding is used in Section 2.11.

Remark 2.2 (Mathematicausage). The `LieTree` algorithm is provided by the script in Table 2.2. Copy it into a text file "TreeComp" in your working directory; to use it, begin your own Mathematicanotebook with the line
`«TreeComp;`

For any symbolic (as distinguished from numerical or graphical) computation, the entries of a real matrix must be in the field $\mathbb{Q}$ (ratios of integers, *not* decimal fractions!) or an algebraic extension of it by algebraic numbers. For instance, if the field is $\mathbb{Q}(\sqrt{-2})$ then Mathematicawill find two linear factors of the polynomial $x^2 + 2y^2$. In symbolic computation, decimal fractions are treated as inexact quantities.

In Mathematica, a matrix B is stored as a list of rows; `Flatten[B]` concatenates the rows. The operation $b = B^\flat$ taking a matrix to a column vector (called vec(B) by Horn and Johnson [132]) defined by (A.8) is implemented in Mathematica as `b=Flatten[Transpose[B]]`. Its inverse $\mathbb{R}^\nu \to \mathbb{R}^{n\times n} : b \mapsto b^\sharp$ is $B :=$ `Transpose[Partition[b,n]]`. △

2.3 Lie Groups

A Lie group is a group and at the same time a manifold on which the group operation and inverse are real-analytic, as described in Appendix B, Sections B.4–B.5.3. After the basic information in Section 2.3.1 and some preparation in Section 2.3.2, Section 2.3.3 attempts to make it plausible that the transition group $\mathbf{G} = \boldsymbol{\Phi}(B_1, \dots, B_m)$ of (2.1) can be given a real-analytic atlas and is indeed a Lie group. The main result of Section 2.3.4 shows that every element of $\mathbf{G}$ is the product of exponentials of the m generators. Section 2.3.5 concerns conditions under which a subgroup of a matrix Lie group is a Lie subgroup.

Example 2.4 (Some Lie groups).
1. The linear space $\mathbb{R}^n$ with composition $(x, y) \mapsto x + y$ (obviously real-analytic) and identity element $\iota = 0$ is an Abelian Lie group. The constant vector fields $\dot{x} = a$ are both right-invariant and left-invariant on $\mathbb{R}^n$ (see page 240).
2. The torus groups $T^n \simeq \mathbb{R}^n/\mathbb{Z}^n$ look like $\mathbb{R}^n$ locally but require several overlapping coordinate[9] charts; for instance, $T^1 = U_1 \bigcup U_2$ where $U_1 = \{\theta \,|\, |\theta| < 2/3\}$, $U_2 = \{\theta \,|\, |\theta - 1| < 2/3\}$.
3. The non-zero complex numbers $\mathbb{C}_*$, with multiplication as the group operation and $\iota = 1$, can be equipped in a $\mathcal{C}^\omega$ way with polar coordinates. The Lie group $\mathbb{C}_*$ is isomorphic to a subgroup of GL$(2, \mathbb{R})$ called the alpha representation of $\mathbb{C}_*$; its Lie algebra $\mathbb{C}$ is represented in this way.

$$\boldsymbol{\alpha}(\mathbb{C}_*) := \left\{ \begin{bmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{bmatrix}, \ a \in \mathbb{R}^2_* \right\}; \quad \boldsymbol{\alpha}(\mathbb{C}) := \left\{ \begin{bmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{bmatrix}, \ a \in \mathbb{R}^2 \right\}. \tag{2.14}$$

2.3.1 *Matrix Lie Groups*

The set GL$(n, \mathbb{R})$ of invertible real matrices $X = \{x_{i,j}\}$ is an open subset of the linear space $\mathbb{R}^{n\times n}$; it inherits the topology and metric of $\mathbb{R}^{n\times n}$, and it is a group under matrix multiplication with identity $\iota = I$. The entries $x_{i,j}$ are real-analytic functions on GL$(n, \mathbb{R})$; the group operation, matrix multiplication, is polynomial in the entries and matrix inversion is rational, so both operations are real-analytic. We conclude that GL$(n, \mathbb{R})$ meets the definition of a Lie group. It has two components separated by the set of singular matrices $\Delta^{(n)}$.

[9] See Definition B.2.

The connected component of $\mathrm{GL}(n, \mathbb{R})$ containing I is a normal subgroup, denoted by $\mathrm{GL}^+(n, \mathbb{R})$ because the determinants of its elements are positive.

Remark 2.3. The vector field $\mathbf{a}$ on $\mathrm{GL}^+(n, \mathbb{R})$ corresponding to $\dot{X} = AX$ is right-invariant on $\mathrm{GL}^+(n, \mathbb{R})$ (see page 240). That follows from the fact that any $S \in \mathrm{GL}^+(n, \mathbb{R})$ defines for each t a right translation $Y(t) = R_S(X(t)) := X(t)S$; it satisfies $\dot{Y}(t) = AY(t)$.

There are other useful types of vector field on $\mathrm{GL}(n, \mathbb{R})$, still linear, such as those given by $\dot{X} = A^\top X + XA$ or $\dot{X} = AX - XA$; these are neither right- nor left-invariant. △

Definition 2.2 (Matrix Lie group). If a matrix group $\mathbf{G}$ is a Lie subgroup of $\mathrm{GL}^+(n, \mathbb{R})$ (simultaneously a submanifold and a subgroup, see Section B.5.3) then $\mathbf{G}$ will be called a matrix Lie group (A matrix Lie group is connected; being a manifold, it can be covered by coordinate charts; see Section B.3.1). △

As an alternative to Definition B.13, the Lie algebra of $\mathbf{G} \subset \mathrm{GL}^+(n, \mathbb{R})$ can be defined[10] as

$$\mathfrak{g} := \{X \in \mathbb{R}^{n\times n} \mid \exp tX \in \mathbf{G} \text{ for all } t \in \mathbb{R}\}.$$

For example, $\{Z \mid \exp(tZ) \in \mathrm{GL}^+(n, \mathbb{R})\} = \mathfrak{gl}(n, \mathbb{R})$. As Lie algebras, $\mathcal{L}(\mathbf{G})$ and $\mathfrak{g}$ are isomorphic: $\mathbf{a} = x^\top A^\top \partial/\partial x$ where $\mathbf{a} \in \mathcal{L}(\mathbf{G})$ and $A \in \mathfrak{g}$.

2.3.2 *Preliminary Remarks*

The set of transition matrices $\boldsymbol{\Phi}(\mathbf{B}^m)$ for (2.2) is a matrix group (subgroup of $\mathrm{GL}^+(n, \mathbb{R})$) as shown in Proposition 1.7. Section 2.3.3 will argue that it is a matrix Lie group, but some preparations are needed.

To see how $\boldsymbol{\Phi}(\mathbf{B}^m)$ and the Lie algebra $\mathbf{B}^m_{\mathcal{L}}$ fit into the present picture, provide every $B_i \in \mathbf{B}^\ell$ with an input u_i to obtain the completed systems

$$\dot{X} = \sum_{i=1}^{\ell} u_i B_i X, \;\; X(0) = I, \tag{2.15}$$

$$\dot{x} = \sum_{i=1}^{\ell} u_i B_i x, \;\; x(0) = \xi. \tag{2.16}$$

Now our set of control matrices $\mathbf{B}^\ell$ is real-involutive, so $\mathrm{span}(\mathbf{B}^\ell)$ is a Lie algebra.

[10] If $\mathbf{G}$ is a *complex Lie group* then [282, Th. 2.10.3(2)] its Lie algebra is $\mathfrak{g} := \{X \mid \exp(sX) \in \mathbf{G}, s \in \mathbb{C}\}$.

Property $\mathfrak{p}$:	$\mathbf{G}_{\mathfrak{p}} \subset \mathrm{GL}^+(n,\mathbb{R})$:
(a) $x_{i,j} = 0,\ i \neq j,$	$\mathrm{D}^+(n,\mathbb{R})$
(b) $X^\top X = I$	$\mathrm{SO}(n)$
(c) $x_{i,j} = \delta_{i,j}$ for $i \leq j$	$\mathrm{N}(n,\mathbb{R})$
(d) $x_{i,j} = 0$ for $i < j$	$\mathfrak{T}^u(n,\mathbb{R})$
(e) $\det(X) = 1$	$\mathrm{SL}(n,\mathbb{R})$
(f) $XH = H$	$\mathrm{Fix}(H)$

Table 2.3. Some well-known closed subgroups of $\mathrm{GL}^+(n,\mathbb{R})$; compare Table 2.1. H is a given matrix, $\operatorname{rank} H \leq n$.

If span $\mathbf{B}^\ell$ has one of the properties and dimensions given in Table 2.1, it can be identified as a Lie subalgebra of $\mathfrak{gl}(n,\mathbb{R})$. Then its corresponding connected Lie group is in Table 2.3 below and has lots of good properties. In general, however, we need to show that the matrix group $\boldsymbol{\Phi}(\mathbf{B}^m)$ can be given a compatible manifold structure (atlas) $\mathbf{U}_A$ and thus qualify as a Lie group; a way of doing that will be sketched in Section 2.3.3.

In Section B.5.3, *submanifold* means, as in Definition B.4, an immersed submanifold; it is not necessarily embedded nor topologically closed in $\mathbf{G}$, as in Example 2.5. In orthodox Lie theory, the assumption is usually made that all subgroups in sight are closed. If that were possible here, the one-to-one correspondence, via the exponential map, between a Lie algebra $\mathfrak{g}$ and a corresponding connected Lie group $\mathbf{G}$ would be an immediate corollary of Theorem B.10. That approach to Lie group theory usually uses the machinery of differentiable manifold theory. To avoid that machinery, one can develop the Lie algebra–Lie group correspondence specialized to matrix Lie groups using the Campbell–Baker–Hausdorff formula.[11] The following proposition and example should be kept in mind as the argument evolves. (The torus T^k is the direct product of k circles.) See Conlon [64, Exer.5.3.8].

Proposition 2.5 (L. Kronecker). *If the set of real numbers $\omega_1, \ldots, \omega_k$, $k \geq 2$, is linearly independent over $\mathbb{Q}$ then the orbits of the real dynamical system $\dot{\theta}_1 = \omega_1, \ldots, \dot{\theta}_k = \omega_k$ are dense in $T^k \simeq \mathbb{R}^k/\mathbb{Z}_k$.*

Example 2.5. A dynamical system of two uncoupled simple harmonic oscillators (think of two clocks with different rates) has the dynamics

$$\dot{x} = Fx, \text{ where } F = \begin{bmatrix} J & 0_{2,2} \\ 0_{2,2} & \omega J \end{bmatrix} \in \mathbb{R}^{4\times 4},$$

[11] Hall [115] (for closed Lie subgroups) and Hilgert et al. [127] (for Lie semigroups) have used the CBH formula as the starting point to establish manifold structure. For a good exposition of this approach to matrix Lie groups (not necessarily closed), see Rossman's book [225].

whose solutions for $t \in \mathbb{R}$ are actions on $\mathbb{R}^4$ of the Lie group

$$\mathbf{G} := e^{\mathbb{R}F} \subset T^2 \subset \mathrm{GL}^+(4, \mathbb{R});$$
$$\mathbf{G} = \left\{ \begin{bmatrix} \cos(t) & \sin(t) & 0 & 0 \\ -\sin(t) & \cos(t) & 0 & 0 \\ 0 & 0 & \cos(\omega t) & \sin(\omega t) \\ 0 & 0 & -\sin(\omega t) & \cos(\omega t) \end{bmatrix}, \ t \in \mathbb{R} \right\}.$$

If $\omega \in \mathbb{Q}$, then $\mathbf{G}$ is a closed subgroup whose orbit is a closed curve on the 2-torus T^2. If ω is *irrational*, $\mathbf{G}$ winds around the torus infinitely many times; in that case, giving it the intrinsic topology of $\mathbb{R}$, $\mathbf{G}$ is a real-analytic manifold immersed in the torus. Note that it is the limit of the compact submanifolds $\{e^{tF} \mid |t| \le T\}$. However, the closure of $\mathbf{G}$ is T^2; $\mathbf{G}$ is dense and the integral curve e^{tF} comes as close as you like to each $p \in T^2$. (Our two clocks, both started at midnight on Day 0, will never again exactly agree.) △

In the next section, it will be useful to assume that the basis for $\mathfrak{g} = \mathrm{span}(\mathbf{B}^\ell)$ is *orthonormal* with respect to the inner product and Frobenius norm on $\mathfrak{gl}(n, \mathbb{R})$ given, respectively, by

$$\langle X, Y \rangle = \mathrm{tr}(X^\top Y), \ \|X\|_2^2 = \langle X, X \rangle.$$

Using the Gram–Schmidt method, we can choose real linear combinations of the original B_i to obtain such a basis and relabel it as $\{B_1, \dots, B_\ell\}$ with $\langle B_i, B_j \rangle = \delta_{i,j}$. Complete this basis, again with Gram–Schmidt, to an orthonormal basis for $\mathfrak{gl}(n, \mathbb{R})$, $\{B_1, \dots, B_\ell, B_{\ell+1}, \dots, B_\nu\}$. The span of the $\nu - \ell$ matrices orthogonal to $\mathfrak{g}$ is called $\mathfrak{g}^\perp$, the orthogonal complement of $\mathfrak{g}$, and is not necessarily a Lie algebra.

Exercise 2.2. If you write $X \in \mathfrak{gl}(n, \mathbb{R})$ as a vector $X^\flat \in \mathbb{R}^\nu$, then the above inner product corresponds to the Euclidean one:

$$\langle X, Y \rangle = \sum_{k=1}^{\nu} X_k^\flat Y_k^\flat. \qquad \triangle$$

2.3.3 *To Make a Manifold*

To show that $\mathbf{G} := \boldsymbol{\Phi}(\mathbf{B}^\ell)$ is a Lie group, it must be shown that it can be provided with a real-analytic atlas $\mathbf{U}_A$. That could be done with a method that uses Theorem B.10, which is based on Frobenius' Theorem and maximal integral manifolds.[12] The following statement and sketched argument attempt to avoid that method without assuming that $\mathbf{G}$ is closed in $\mathrm{GL}^+(n, \mathbb{R})$.

[12] See Varadarajan [282, Th. 2.5.2].

Proposition 2.6. *The transition-matrix group* $\mathbf{G} := \boldsymbol{\Phi}(\mathbf{B}^\ell)$ *can be provided with a real-analytic atlas* $\mathbf{U}_A$*; thus,* $\mathbf{G}$ *is a Lie group. Its Lie algebra* $\mathcal{L}(\mathbf{G})$ *is* $\mathfrak{g} = \mathrm{span}(\mathbf{B}^\ell)$.

Proof (Sketch).
The attainable sets $\{\mathcal{A}_T(I), T > 0\}$ for (2.2) will serve as a neighborhood base for $\boldsymbol{\Phi}$ as a topological group; see Section B.4.1. Using our orthonormal basis, the matrices $Z(u) = \sum_1^\ell u_i B_i$ will satisfy

$$u_i = \langle Z(u), B_i \rangle,\ 1 \le i \le \ell.$$

The map $Z \to \exp(Z) := I + Z + \cdots$ is real-analytic; therefore so is $u \to \exp(Z(u)) \in \mathbf{G}$. Let

$$J_i(u) := \frac{\partial \mathrm{e}^{Z(u)}}{\partial u_i};\ \text{since } \langle J_i(0), B_j \rangle = \delta_{i,j},\ 1 \le i, j \le \ell,$$

the exponential map on $\mathfrak{g}$ has rank ℓ near $u = 0$. For any $\delta > 0$, let

$$U_\delta := \left\{ u \Big| \sum_1^\ell u_i^2 < \delta^2 \right\},\ Z(U_\delta) := \{ Z(u) \,|\, u \in U_\delta \} \subset \mathfrak{g}.$$

If $u \in U_\delta$, by orthonormality $\|Z(u)\|_2 < \delta$. Let $V_\delta := \exp(Z(U_\delta))$; then each V_δ is a neighborhood of $I \in \mathbf{G}$. Fix some $\delta < \log(2)$; then we can compute the u coordinates of any $Q \in V_\delta$ by

$$u_i = \langle B_i, \log(Q) \rangle \tag{2.17}$$

using the series (1.13) for $\log(Q)$, convergent in the ball $\{Q \,|\, \|Q - I\| < 1\}$. The coordinate functions $\mathbf{u} : V_\delta \to (u_1, \ldots, u_\ell)$ given by (2.17) are called canonical coordinates of the first kind.[13] This topic is continued in Section 2.8.

Choose a matrix $C \in V_\delta$ and a matrix $A \in \mathfrak{g}$, let $Z = \log(C)$, and construct the integral curve of $\dot{X} = AX$ with $X(0) = C$, which is $X(t) = \mathrm{e}^{tA}\mathrm{e}^{Z}$, and (because $\mathfrak{g}$ is a Lie algebra) apply the Campbell–Baker–Hausdorff theorem and its corollary (2.12) to see that for sufficiently small $|t|$, $\log(X(t)) \in Z(U_\delta)$ and thus $X(t) \in V_\delta$. We choose the pair $(V_\delta, \mathbf{u})$ as a fundamental chart for our atlas $\mathbf{U}_A$.

We can cover $\mathbf{G}$ with translates of our fundamental chart, in the following way. For any $T > 0$, the subset

$$\mathbf{G}[T] := \{ X(t, u) \,|\, |u_i| \le 1,\ i \in 1 \,..\, \ell,\ t \in [0, T] \}$$

is compact. There exists a countable set of points $\{\Psi_1, \Psi_2, \ldots\}$ in $\mathbf{G}[T]$ such that $\mathbf{G}[T]$ is covered by the union of the translates $V_\delta \Psi_i$. Letting $T \to \infty$, one sees that the same is true of the topological group $\mathbf{G}$. Define the coordinates

[13] See Helgason [120] and an often-cited work of Magnus [195], which made extensive use of canonical coordinates of the first kind in computational physics.

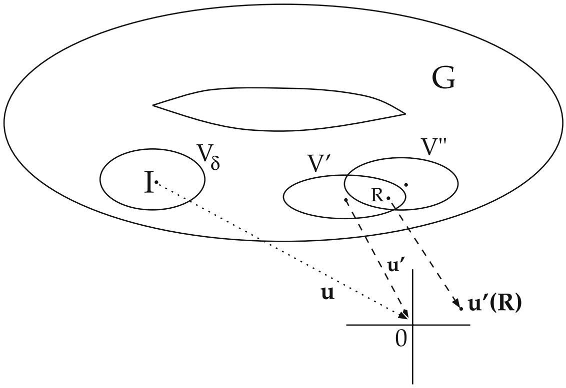


Fig. 2.1. Coordinate charts on **G**.

in each of the translated neighborhoods $V_\delta \Psi_j$ by translating the basic chart: each point $Q \in V_\delta \Psi_j$ has local coordinates

$$\mathbf{u}(Q) = \left\{ \langle B_i, \log(Q\Psi_j^{-1}) \rangle, i \in 1 \ldots \ell \right\}.$$

To obtain the required $\mathbb{C}^\omega$ coordinate transformations between arbitrary charts, it suffices (since **G** is path-connected) to show that in $\mathbf{U}_A$ two sufficiently small overlapping neighborhoods $V' := V_\delta \Psi_1$, $V'' := V_\delta \Psi_2$ in **G** with intersection $V' \cap V'' \neq \varnothing$ can be given respective charts having a $\mathbb{C}^\omega$ coordinate change on $V' \cap V''$.

To build the chart on V'', start with the existence of $Y \in U_\delta$ such that $\Psi_2 = e^Y \Psi_1$ and right-translate V_δ with Ψ_2. Since the charts overlap, there exists $R \in V' \cap V''$ as in Fig. 2.1.

There exist in U_δ matrices $W' := \log(R\Psi_1^{-1})$ and $W'' := \log(R\Psi_2^{-1})$. In V', V'' the coordinates of R are, respectively, $u_i'(R) = \langle W', B_i \rangle$ and $u_i'(R) = \langle W'', B_i \rangle$. That is, with respect to the given basis, W' and W'' can be identified with coordinate vectors at R with respect to the two charts.

The coordinate transformations can be found using the Campbell–Baker–Hausdorff function μ of (2.10):

$$\begin{aligned} W'' &= \log(R\Psi_2^{-1}) = \log(e^{W'} \Psi_1 \Psi_2^{-1}) \\ &= \log(e^{W'} e^{-Y}) = \mu(W', -Y), \text{ so} \\ u_i''(R) &= \langle B_i, \mu(W', -Y) \rangle. \end{aligned} \tag{2.18}$$

$$\begin{aligned} W' &= \log(R\Psi_1^{-1}) = \log(e^{W''} \Psi_2 \Psi_1{}^{-1}) \\ &= \log(e^{W''} e^{Y}) = \mu(W'', Y), \text{ so} \\ u_i' &= \langle B_i, \mu(W'', Y) \rangle. \end{aligned} \tag{2.19}$$

These relations, as the compositions of real-analytic mappings, are real-analytic. In (2.18)–(2.19) remember that $u \in \mathbb{R}^\ell$ corresponds to $W \in \mathfrak{g}$.

Since matrix multiplication is real-analytic, this local chart extends to a real-analytic immersion $\mathbf{G} \hookrightarrow GL^+(n, \mathbb{R})$. Suppose that $\mathbf{G}$ intersects itself at some point S. By the definition of $\mathbf{G}$, there are some matrices $Y_1, \dots, Y_j \in \mathfrak{g}$, $Z_1, \dots, Z_k \in \mathfrak{g}$ such that the two transition matrices

$$F := \exp(Y_k) \cdots \exp(Y_2) \exp(Y_1),\ F' := \exp(Z_k) \cdots \exp(Z_2) \exp(Z_1)$$

satisfy $F = S$ and $F' = S$. F and F' have positive determinants. Given U, a sufficiently small neighborhood of I in $\mathbf{G}$, FU and $F'U$ are neighborhoods of Q in $\mathbf{G}$, and each takes $\mathfrak{g}$ to itself, preserving orientation. By the uniqueness of the trajectories on $GL^+(n, \mathbb{R})$ of $\dot{X} = AX$, $A \in \mathfrak{g}$, $X(0) = S$, we see that there is a neighborhood of I in $FU \cap F'U$. That is, self-intersections are local inclusions.

Return to the chart at the identity in $GL^+(n, \mathbb{R})$ given by

$$\log : e^{u_1 B_1 + \cdots + u_\nu B_\nu} \mapsto u \in \mathbb{R}^\nu;$$

provided that ϵ is sufficiently small the set

$$S_\epsilon := \{\exp(tE) | -1 < t < 1,\ E \in \mathfrak{g}^\perp, \|E\| < \epsilon\}$$

constructed at $I \in GL^+(n, \mathbb{R})$ can be assigned the $\nu - \ell$ coordinates obtained by

$$u_j = \langle Z(u), B_i \rangle,\ \ell < i \le \nu.$$

S_ϵ is transverse to V_δ, and $S_\epsilon \times V_\delta$ is C^ω-diffeomorphic to a box in $\mathbb{R}^{\nu-\ell} \times \mathbb{R}^\ell$. (It would be nice to say something like this for any right translation of $S_\epsilon \times V_\delta$ by an element $Q \in \mathbf{G}$, $S_\epsilon Q \times V_\delta Q$, but self-intersections of such sets would occur in cases like Example 2.5 no matter how small ϵ may be, unless $\mathbf{G}$ is closed.) This concludes our discussion of the manifold structure of $\mathbf{G}$.

For a vector field $\mathbf{b}$ to be locally right-invariant on $V_\delta \subset \mathbf{G}$, it must be of the form $\dot{X} = BX$ for some $B \in \mathfrak{g}$, which shows the desired correspondence $\mathcal{L}(\mathbf{G}) = \mathfrak{g}$. □

Remark 2.4. The method used here shows that the class of piecewise constant controls $\mathcal{PK}$ is sufficiently large to give as transition matrices all the matrices in $\mathbf{G}$. Nothing new would be obtained by using $\mathcal{LI}$ (locally integrable) controls. △

Another way of charting $\mathbf{G}$ uses the fact that the map

$$(u_1, u_2, \dots, u_\ell) \to e^{u_\ell B_\ell} \cdots e^{u_2 B_2} e^{u_1 B_1} \in \mathbf{G} \tag{2.20}$$

has rank ℓ at $u = 0$ and is real-analytic. Its inverse provides canonical coordinates of the second kind in V_δ. For more about both types of canonical coordinates, see Section 2.8 and Varadarajan [282, Sec. 2.10].

2.3.4 Exponentials of Generators Suffice

Now return to a general symmetric matrix control system (2.2), whose generating set $\mathbf{B}^m = \{B_1, \dots, B_m\}$ is not necessarily real-involutive but generates a Lie algebra $\mathfrak{g} = \mathbf{B}^m_{\mathcal{L}}$ whose basis is $\mathbf{B}^\ell \subset \mathbf{B}^\nu$. As before, let $\boldsymbol{\Phi}(\mathbf{B}^m)$ denote the group of transition matrices of (2.2). Proposition 2.6 says that $\mathbf{G} := \boldsymbol{\Phi}(\mathbf{B}^\ell)$ is the Lie subgroup of $\mathrm{GL}^+(n, \mathbb{R})$ corresponding to our Lie algebra $\mathfrak{g}$. Here we show in detail how (2.2) gives rise to the same Lie group $\mathbf{G}$ as its completion (2.15); $\mathbf{B}^m$ generates the ℓ-dimensional Lie algebra $\mathfrak{g}$ and the exponentials $\exp(tB_i)$ generate the Lie group $\mathbf{G}$.

Proposition 2.7. $\boldsymbol{\Phi}(\mathbf{B}^m) = \boldsymbol{\Phi}(\mathbf{B}^\ell) = \mathbf{G}$.

Proof. The relation $\boldsymbol{\Phi}(\mathbf{B}^m) \subset \boldsymbol{\Phi}(\mathbf{B}^\ell) = \mathbf{G}$ is obvious. We need to show that any element of $\mathbf{G}$ is in $\boldsymbol{\Phi}(\mathbf{B}^m)$.

If $H_0 := \mathrm{span}\{B_1, \dots, B_m\}$ is real-involutive, then $H_0 = \mathfrak{g}$, $m = \ell$ and we are done.

Otherwise there are elements $B_i, B_j \in H_0$ for which $[B_i, B_j] \notin H_0$, so one can choose τ in an interval $0 < t_1 < \tau$ (depending on B_i and B_j) such that by (2.9)

$$C := e^{t_1 B_i} B_j e^{-t_1 B_i} = B_j + t_1 [B_i, B_j] + \frac{t_1^2}{2}[B_i, [B_i, B_j]] + \cdots$$

is in $\mathfrak{g}$ but not in H_0. Let $B_{m+1} := C$ and let $H_1 = \mathrm{span}\{B_1, \dots, B_{m+1}\}$. Using $\exp(tPAP^{-1}) = Pe^{tA}P^{-1}$,

$$e^{sC} = e^{t_1 B_i} e^{sB_j} e^{-t_1 B_i}$$

and therefore e^{sC} can be reached from I by the control sequence that assigns $\dot{X}$ the value $-B_i$ for t_1 seconds, B_j for s seconds, and B_i for t_1 seconds. If H_1 is real-involutive we are finished. Otherwise continue; there are several possibilities of pairs in H_1 whose bracket might be new, but one will suffice. Suppose $B_k, C_1 \in H_1$ with $[B_k, C_1] \notin H_1$. Then there exist values of t_2 such that

$$\begin{aligned} C_2 :=& e^{t_2 B_k} C_1 e^{-t_2 B_k} \notin H_1; \\ e^{sC_2} =& e^{t_2 B_k} e^{sC_1} e^{-t_2 B_k} = e^{t_2 B_k} e^{t_1 B_i} e^{sB_j} e^{-t_1 B_i} e^{-t_2 B_k} \\ =& e^{t_2 \,\mathrm{ad}_{B_k}} (e^{t_1 \,\mathrm{ad}_{B_j}} (e^{sB_j})). \end{aligned} \tag{2.21}$$

Repeat the above process using $\exp(tB_i)$ operations until we exhaust all the possibilities for new Lie brackets C_i in at most $\ell - m$ steps. (Along the way we obtain a basis $\{B_1, \dots, B_m, C_1, \dots\}$ for $\mathfrak{g}$ that depends on our choices; but any two such bases are conjugate.)

From (2.20), each $Q \in \mathbf{G}$ is the product of a finite number of exponentials $\exp(tB_i)$ of matrices in $\mathbf{B}^\ell$. By our construction above, $\exp(tB_i) \in \boldsymbol{\Phi}(\mathbf{B}^m)$ for $i \in 1, \dots, \ell$, which completes the proof.

Also note that each $Q \in \mathbf{G}$ can be written as a product in which the jth factor is of the form

$$\exp\Big(\tau_j u_{i_j}(\tau_j) B_{i_j}\Big),\ i_j \in \{1, \dots m\},\ \tau_j > 0,\ u_{i_j}(\tau_j) \in \{-1, 1\}.$$

$\square$

Corollary 2.1. *Using unrestricted controls in the function space $\mathcal{LI}$,* (2.1) *generates the same Lie group $\boldsymbol{\Phi}(\mathbf{B}^m) = \mathbf{G}$ that has been obtained in Proposition 2.7 (in which $u \in \mathcal{PK}$ and $|u_i(t)| = 1, i \in 1\,..\,m$).*

Proposition 2.7 is a bilinear version, first given in an unpublished report Elliott and Tarn [85], of a theorem of W. L. Chow [59] (also [258]) fundamental to control theory. Also see Corollary 2.2 below. It should be emphasized (a) that Proposition 2.7 makes the controllability of (2.1) equivalent to the controllability of (2.16) and (b) that although $\boldsymbol{\Phi}(\mathbf{B}^m)$ and $\mathbf{G} = \boldsymbol{\Phi}(\mathbf{B}^\ell)$ are the same as groups, they are generated differently and the control systems that give rise to them are different in important ways. That will be relevant in Section 6.4.

Let $Y(t; \tilde{u})$ be a solution of the m-input matrix control system

$$\dot{Y} = \sum_{i=1}^{m} \tilde{u}_i B_i Y,\ Y(0) = I,\ \Omega := \mathbb{R}^m. \tag{2.22}$$

Consider a completion (orthonormal) of this system but with constrained controls

$$\dot{X} = \sum_{i=1}^{\ell} u_i B_i X,\ X(0) = I,\ u \in \mathcal{LI},\ \|u(t)\| \le 1. \tag{2.23}$$

We have seen that the system (2.22) with m *unconstrained* controls can approximate the missing Lie brackets $B_{m+1}, \dots, B_\ell$ (compare Example 2.1). An early (1970) and important work on Chow's theorem, Haynes and Hermes [119], has an approximation theorem for C^∞ nonlinear systems that can be paraphrased for bilinear systems in the following way.

Proposition 2.8. *Given $\xi \in \mathbb{R}^n_*$, any control u such that $\|u(t)\| \le 1$, the solution $X(t; u)$ of the completed system* (2.23) *on $[0, T]$, and any $\varepsilon > 0$, there exists a measurable control $\tilde{u}$ such that the corresponding solution $Y(t; \tilde{u})$ of* (2.22) *satisfies*

$$\max_{0 \le t \le T} \left\{ \frac{\|Y(t; \tilde{u})\xi - X(t; u)\xi\|}{\|\xi\|} \right\} < \varepsilon. \qquad \triangle$$

2.3.5 Lie Subgroups

Before discussing closed subgroups two points should be noted. The first is to say what is meant by the intrinsic topology $\tau(\mathbf{G})$ of a Lie group $\mathbf{G}$ with Lie algebra $\mathfrak{g}$; as in the argument after Proposition 2.6, $\tau(\mathbf{G})$ can be given by a neighborhood basis obtained by right-translating small neighborhoods of the identity $\{\exp(Z)|\, Z \in \mathfrak{g}, \|Z\| < \epsilon\}$. Second, $\mathbf{H} \subset \mathbf{G}$ is called a closed subgroup of $\mathbf{G}$ if it is closed in $\tau(\mathbf{G})$.

Proposition 2.9. *Suppose that* $\mathbf{G}$ *is a connected subgroup of* $\mathrm{GL}^+(n, \mathbb{R})$. *If*

$$X_i \in \mathbf{G},\ Q \in \mathrm{GL}^+(n, \mathbb{R}),\ \ \textit{and}\ \lim_{i\to\infty} \|X_i - Q\| \to 0\ \textit{imply}\ Q \in \mathbf{G},$$

then $\mathbf{G}$ *is a closed matrix Lie group.*

Proof. The group $\mathbf{G}$ is a matrix Lie group (Definition 2.2). The topology of $\mathrm{GL}^+(n, \mathbb{R})$ is compatible with convergence in matrix norm. In this topology, $\mathbf{G}$ is a closed subgroup of $\mathrm{GL}^+(n, \mathbb{R})$, and it is a regularly embedded submanifold of $\mathrm{GL}^+(n, \mathbb{R})$. □

If a matrix Lie group $\mathbf{G}$ is compact, like $\mathrm{SO}(n)$, it is closed. If $\mathbf{G}$ is not compact, its closure may have dimension greater than $\dim \mathbf{G}$ as in the one-parameter groups of Example 2.5. The next proposition has its motivation in the possibility of such dense subgroups. As for the hypothesis, all closed subgroups of $\mathrm{GL}(n, \mathbb{R})$ are locally compact. For a proof of Proposition 2.10, see Hochschild [129, XVI] (Proposition 2.3). Rather than "connected Lie group," Hochschild, following Chevalley [58], uses the term analytic group.

Proposition 2.10. *Let* γ *be a continuous homomorphism of* $\mathbb{R}$ *into a locally compact group* $\mathbf{G}$. *Then either* $\gamma \to \mathbb{R}$ *is onto* $\gamma(\mathbb{R})$ *or the closure of* $\gamma(\mathbb{R})$ *in* $\mathbf{G}$ *is compact.*

There is an obvious need for conditions sufficient for a subgroups to be closed. Two of them are easy to state. The first is the criterion of Malcev [196].[14]

Theorem 2.2 (Malcev). *Let* $\mathbf{H}$ *be a connected Lie subgroup of a connected Lie group* $\mathbf{G}$. *Suppose that the closure in* $\mathbf{G}$ *of every one-parameter Lie group of* $\mathbf{H}$ *lies in* $\mathbf{H}$. *Then* $\mathbf{H}$ *is closed in* $\mathbf{G}$.

The second criterion employs a little algebraic geometry. A real algebraic group $\mathbf{G}$ is a subgroup of $\mathrm{GL}(n, \mathbb{R})$ that is also[15] the zero-set of a finite collection of polynomial functions $p : \mathbb{R}^{n\times n} \to \mathbb{R}$. See Section B.5.4: by Theorem B.7, $\mathbf{G}$ is a closed (Lie) subgroup of $\mathrm{GL}(n, \mathbb{R})$.

The groups $\mathbf{G}_a$–$\mathbf{G}_f \subset \mathrm{GL}^+(n, \mathbb{R})$ in Table 2.3 are defined by polynomial equations in the matrix elements, such as $XX^\top = I$, that are invariant under matrix multiplication and inversion, so they are algebraic groups. We will refer to them often. As Lie groups, $\mathbf{G}_a$–$\mathbf{G}_f$ respectively correspond to the Lie algebras of Table 2.1. To generalize Property (b) let Q be symmetric, nonsingular, with signature $\{p, q\}$; to the Lie algebra $\mathfrak{so}(p, q)$ of (2.5) corresponds the Lie group $\mathrm{SO}(p, q) := \{X|\, X^\top QX = Q\}$.

[14] For Theorem 2.2, see the statement and proof in Hochschild [129, XVI, Proposition 2.4] instead.

[15] In the language of Appendix C, an algebraic group is an affine variety in $\mathbb{R}^{n\times n}$.

2.4 Orbits, Transitivity, and Lie Rank

Proposition 2.7 implies that the attainable set through a given $\xi \in \mathbb{R}^n$ for either (2.1) or (2.16) is the orbit $\mathbf{G}\xi$. (See Example 2.12.) Once the Lie algebra $\mathfrak{g}$ generated by $\mathbf{B}^m$ has been identified, the group $\mathbf{G}$ of transition matrices is known in principle as a Lie subgroup, not necessarily closed, of $GL^+(n, \mathbb{R})$. The theory of controllability of symmetric bilinear systems is thus the study of the orbits of $\mathbf{G}$ on $\mathbb{R}^n$, which will be the chief topic of the rest of this chapter and lead to algebraic criteria for transitivity on $\mathbb{R}^n_*$. Section 2.5.1 shows how orbits may be found and transitivity decided.

Controllability of (2.1) is thus equivalent to transitivity of $\mathbf{G}$. That leads to two questions:

Question 2.1. Which subgroups $\mathbf{G} \subset GL^+(n, \mathbb{R})$ are transitive on $\mathbb{R}^n_*$?

Question 2.2. If $\mathbf{G}$ is not transitive on $\mathbb{R}^n$, what information can be obtained about its orbits?

Question 2.2 will be examined beginning with Section 2.7. In textbook treatments of Lie theory, one is customarily presented with some named and well-studied Lie algebra (such as those in Table 2.1) but in control problems we must work with whatever generating matrices are given.

Question 2.1 was answered in part by Boothby [29] and Boothby and Wilson [32], and completed recently by Kramer [162]. Appendix D lists all the Lie algebras whose corresponding subgroups of $GL^+(n, \mathbb{R})$ act transitively on $\mathbb{R}^n_*$. The Lie algebra $\mathfrak{g}$ corresponding to such a transitive group is called a transitive Lie algebra; it is itself a transitive set, because $\mathfrak{g}x = \mathbb{R}^n$ for all nonzero x, so the word "transitive" has additional force. We will supply some algorithms in Section 2.5.2 that, for low dimensions, find the orbits of a given symmetric system and check it for transitivity.

Example 2.6. For $k = 1, 2$ and 4, among the Lie groups transitive on $\mathbb{R}^k_*$ is the group of invertible elements of a normed division algebra.[16]

A normed division algebra $\mathfrak{D}$ is a linear space over the reals such that there exists

(i) a multiplication $\mathfrak{D} \times \mathfrak{D} \to \mathfrak{D}$: $(a, b) \mapsto ab$, bilinear over $\mathbb{R}$;
(ii) a unit 1 such that for each $a \in \mathfrak{D}$ $1a = a1 = a$;
(iii) a norm $\|\cdot\|$ such that $\|1\| = 1$ and $\|ab\| = \|a\|\,\|b\|$, and
(iv) for each nonzero element a, elements $r, l \in \mathfrak{D}$ such that $la = 1$ and $ar = 1$.
Multiplication need not necessarily be commutative or associative.

It is known that the only normed division algebras $\mathfrak{D}^d$ (where d is the dimension over $\mathbb{R}$) are $\mathfrak{D}^1 = \mathbb{R}$; $\mathfrak{D}^2 = \mathbb{C}$; $\mathfrak{D}^4 = \mathbb{H}$, the quaternions (noncommutative); and $\mathfrak{D}^8 = \mathbb{O}$, the octonions (noncommutative and *nonassociative*

[16] See Baez [16] for a recent and entertaining account of normed division algebras, spinors, and octonions.

multiplication). If $ab \neq 0$, the equation $xa = b$ can be solved for $x \neq 0$, so the subset $\mathfrak{D}^k_* := \mathfrak{D}^k \setminus \{0\}$ acts transitively on itself; the identifications $\mathfrak{D}^k_* \simeq \mathbb{R}^k_*$ show that there is a natural and transitive action of $\mathfrak{D}^k_*$ on $\mathbb{R}^k_*$. $\mathfrak{D}^k_*$ is a manifold with $\mathbb{C}^\omega$ global coordinates. For $k = 1, 2, 4$, $\mathfrak{D}^k_*$ is a topological group; therefore, it is a Lie group and transitive on $\mathbb{R}^k_*$. The Lie algebras of $\mathbb{C}_*$ and $\mathbb{H}_*$ are isomorphic to $\mathbb{C}$ and $\mathbb{H}$; they are isomorphic to the Lie algebras given in Section D.2 as, respectively, I.1 and I.2 (with a one-dimensional center).

Since its multiplication is not associative, $\mathfrak{D}^8_*$ is *not* a group and has no Lie algebra. However, the octonions are used in Section D.2 to construct a Lie algebra for a Lie group **Spin**$(9, 1)$ that is transitive on $\mathbb{R}^{16}_*$. △

2.4.1 *Invariant Functions*

Definition 2.3. A function ϕ is said to be **G**-invariant, or invariant under **G**, if for all $P \in \mathbf{G}$, $\phi(Px) = \phi(x)$. If there is a function $\theta : \mathbf{G} \to \mathbb{R}$ such that $\phi(Px) = \theta(P)\phi(x)$ on $\mathbb{R}^n$, we will say ϕ is relatively **G**-invariant with weight θ. △

Example 2.7 (Waterhouse [283]). Let **G** be the group described by the matrix dynamical system on $\mathbb{R}^{n\times n}$

$$\dot{X} = [A, X], X(0) = C \in \mathbb{R}^{n\times n}, \quad X(t) = e^{tA}Ce^{-tA};$$

it has the obvious invariant $\det(X(t)) = \det(C)$. In fact, for any initial $C \in \mathbb{R}^{n\times n}$ and $\lambda \in \mathbb{C}$, the determinant

$$\det(\lambda I_n - X(t)) = \det\Big(e^{tA}(\lambda I_n - C)e^{-tA}\Big) = \det(\lambda I_n - C) = \sum_{j=0}^{n} p_{n-j}\lambda^j$$

is constant with respect to t, so the coefficient p_{n-j} of λ^j in $\det(\lambda I_n - X(t))$ is constant. That is, $p_0, \dots, p_n$, where $p_0 = 1$ and $p_n = \det(C)$, are invariants under the adjoint action of e^{tA} on C.

This idea suggests a way to construct examples of invariant polynomials on $\mathbb{R}^\nu$. Lift $\dot{X} = [A, X]$ from $\mathbb{R}^{n\times n}$ to a dynamical system on $\mathbb{R}^\nu$ by the one-to-one mapping

$$X \mapsto z = X^\flat \text{ with inverse } X = z^\sharp; \tag{2.24}$$
$$\dot{z} = \tilde{A}z;\ z(0) = C^\flat,\ \text{where } \tilde{A} := I_n \otimes A - A^\top \otimes I_n.$$

Let $q_j(z) := p_j(z^\sharp) = p_j(C)$; then $q_j(z)$, which is a jth degree polynomial in the entries of z, is invariant for the one-parameter subgroup $\exp(\mathbb{R}\tilde{A}) \subset \mathrm{GL}^+(\nu, \mathbb{R})$; that is, $q_j(\exp(t\tilde{A})z) = q_j(z)$.

This construction easily generalizes from dynamical systems to control systems, because the *same* polynomials $p_j(C)$ and corresponding $q_j(z)$ are invariant for the matrix control system

$$\dot{X} = (uA + vB)\,X - X\,(uA + vB),\;\; X(0) = C$$

and for the corresponding bilinear system on $\mathbb{R}^\nu$ given by (2.24). △

To construct more examples, it is convenient to use vector fields and Lie algebras. Given a (linear or nonlinear) vector field $\mathbf{f}$, a function $\psi(x)$ on $\mathbb{R}^n$ is called an invariant function for $\mathbf{f}$ if $\mathbf{f}\psi(x) = 0$. There may also be functions ϕx that are relatively invariant for $\mathbf{f}$; that means, there exists a real $\lambda \neq 0$ depending only on $\mathbf{f}$ for which $\mathbf{f}\phi(x) = \lambda\phi(x)$. Then $\dot{\phi} = \lambda\phi$ along any trajectory of $\mathbf{f}$.

If $\mathbf{f}\psi = \lambda\psi$, $\mathbf{g}\psi = \mu\psi$ and the Lie algebra they generate is $\mathfrak{g} := \{\mathbf{f}, \mathbf{g}\}_{\mathcal{L}}$, then ψ is invariant for the vector fields in the derived Lie subalgebra $\mathfrak{g}' := [\mathfrak{g}, \mathfrak{g}]$. Note in what follows the correspondence between matrices A, B and linear vector fields $\mathbf{a}, \mathbf{b}$ given by (2.4).

Proposition 2.11. [17] *Let $\mathfrak{g} := \{\mathbf{b}_1, \dots, \mathbf{b}_m\}_{\mathcal{L}}$ on $\mathbb{R}^n$. If there exists a real-analytic function ϕ such that $\phi(x(t)) = \phi(x(0))$ along the trajectories of* (2.1) *then*
(i) ϕ must satisfy this eigenvalue problem using the generating vector fields:

$$\mathbf{b}_1\phi = \lambda_1\phi, \dots, \mathbf{b}_m\phi = \lambda_m\phi. \quad (\mathit{If}\ \mathbf{b}_i \in \mathfrak{g}',\ \mathit{then}\ \lambda_i = 0.) \tag{2.25}$$

(ii) The set of functions $\mathfrak{F} := \{\phi |\, \mathbf{c}\phi = 0 \ \text{for all } \mathbf{c} \in \mathfrak{g}\} \subset C^\omega(\mathbb{R}^n)$ is closed under addition, multiplication by reals, multiplication of functions, and composition of functions.

Proof. Use $\dot{\phi} = u_1\mathbf{b}_1\phi + \cdots + u_m\mathbf{b}_m\phi$ for all u and $[\mathbf{a}, \mathbf{b}]\phi = \mathbf{b}(\mathbf{a}\phi) - \mathbf{a}(\mathbf{b}\phi)$ to show (i). For (ii), use the fact that $\mathbf{a}$ is a derivation on $C^\omega(U)$ and that if $\phi \in \mathfrak{F}$ and $\psi \in \mathfrak{F}$ then $\mathbf{f}\phi \circ \psi = \phi_* \mathbf{f}\psi = 0$. □

Definition 2.4. A function $\psi : \mathbb{R}^n \to \mathbb{R}$ is called homogeneous of degree k if $\psi(\lambda x) = \lambda^k \psi(x)$ for all $\lambda \in \mathbb{R}$. △

Example 2.8. Homogeneous differentiable functions are *relatively* invariant for the Euler vector field $\mathbf{e} := x^\top \partial/\partial x$; if ψ is homogeneous of degree r then $\mathbf{e}\psi = r\psi$. The invariant functions of degree zero for $\mathbf{e}$ are independent of $\|x\|$. △

Proposition 2.12. *Given any matrix Lie group* $\mathbf{G}$*, its Lie algebra $\mathfrak{g}$, and a real-analytic function $\phi : \mathbb{R}^n_* \to \mathbb{R}$, the following statements about ϕ are equivalent:*
(a) $\mathbf{a}\phi = 0$ *for all vector fields* $\mathbf{a} \in \mathfrak{g}$.
(b) ϕ is invariant under the action of $\mathbf{G}$ *on $\mathbb{R}^n_*$.*
(c) If $\{B_1, \dots, B_m\}_{\mathcal{L}} = \mathfrak{g}$ then ϕ is invariant along the trajectories of (2.1).

[17] Proposition 2.11 and its proof hold for nonlinear real-analytic vector fields; it comes from the work of George Haynes[118] and Robert Hermann [123] in the 1960s.

Proof. (c) $\Leftrightarrow$ (b) by Proposition 2.7. The statement (a) is equivalent to $\phi(\exp(tC)x) = \phi(x)$ for all $C \in \mathfrak{g}$, which is (b). $\square$

Given any matrix Lie group, what are its invariant functions? Propositions 2.14 and 2.15 will address that question. We begin with some functions of low degree.

All invariant linear functions $\psi(x) = c^\top x$ of $\mathbf{g} := \{\mathbf{b}_1, \dots, \mathbf{b}_m\}_{\mathcal{L}}$ are obtained using (2.25) by finding a common left nullspace

$$N := \{c |\, c^\top B_1 = 0, \dots, c^\top B_m = 0\};$$

then $N\mathbf{G} = N$.

The polynomial $x^\top Qx$ is $\mathbf{G}$-invariant if $B_i^\top Q + QB_i = 0$, $i \in 1, \dots, m$. Linear functions $c^\top x$ or quadratic functions $x^\top Qx$ that are relative $\mathbf{G}$ invariants satisfy an eigenvalue problem, respectively, $c^\top B_i = \lambda_i c^\top$, $B_i^\top Q + QB_i = \mu_i Q$, $i \in 1, \dots, m$. If B_i belongs to the derived subalgebra $\mathfrak{g}'$ then $\lambda_i = 0$ and $\mu_i = 0$.

Example 2.9 (Quadric orbits). The best-known quadric orbits are the spheres in $\mathbb{R}^3_*$, invariant under SO(3). The quadratic forms with signature $\{p, q\}$, $p+q = n$, have canonical form $\psi(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 \cdots - x_n^2$ and are invariant under the Lie group SO(p, q); the case $\{3, 1\}$ is the Lorentz group of special relativity, for which the set $\{x |\, \psi(x) = 0\}$ is the light-cone.

Some of the transitive Lie algebras listed in Appendix D are of the form $\mathfrak{g} = \mathfrak{c} \oplus \mathfrak{g}_1$, where $\mathfrak{g}_1$ is a Lie subalgebra of $\mathfrak{so}(n)$ and

$$\mathfrak{c} := \{X \in \mathfrak{g} |\, \mathrm{ad}_Y(X) = 0 \ \text{ for all } \ Y \in \mathfrak{g}\}$$

is the center of $\mathfrak{g}$. In such cases $x^\top Qx$, where $Q \gg 0$, is an invariant polynomial for $\mathfrak{g}_1$ and a relative invariant for $\mathfrak{c}$. $\triangle$

Example 2.10. The identity $\det(QX) = \det(Q)\det(X)$ shows that the determinant function on the matrix space $\mathbb{R}^{n\times n}$ is SL$(n, \mathbb{R})$-invariant and that its level sets, such as the set $\Delta^{(n)}$ of singular matrices, are invariant under SL$(n, \mathbb{R})$'s action on $\mathbb{R}^{n\times n}$. The set $\Delta^{(n)}$ is interesting. It is a semigroup under multiplication; it is not a submanifold, since it is the disjoint union of SL$(n, \mathbb{R})$-invariant sets of different dimensions. $\triangle$

2.4.2 *Lie Rank*

The idea of Lie rank originated in the 1960s in work of Hermann [122], Haynes [118] (for real-analytic vector fields) and Kučera [167–169] (for bilinear systems).

If the Lie algebra $\mathfrak{g}$ of (2.1) has the basis $\mathbf{B}^\ell = \{B_1, \dots, B_m, \dots, B_\ell\}$, at each state x the linear subspace $\mathfrak{g}x \subset \mathbb{R}^n$ is spanned by the ℓ columns of the matrix

$$M(x) := \left[B_1x\ B_2x\ \dots\ B_mx\ \cdots\ B_\ell x\right]. \tag{2.26}$$

$$\text{Let } \rho_x(\mathfrak{g}) := \operatorname{rank} M(x). \tag{2.27}$$

The integer-valued function ρ, defined for any matrix Lie algebra $\mathfrak{g}$ and state x, is called the Lie rank of $\mathfrak{g}$ at x. The Lie algebra rank condition (abbreviated as LARC) is

$$\rho_x(\mathfrak{g}) = n \qquad \text{for all } x \in \mathbb{R}^n_*. \tag{2.28}$$

The useful matrix $M(x)$ can be called the LARC matrix. If $\operatorname{rank}(M(x)) \equiv n$ on $\mathbb{R}^n_*$ and $\phi \in C^\omega(\mathbb{R}^n)$ is $\mathbf{G}$-invariant then ϕ is constant, because

$$\frac{\partial \phi^\top}{\partial x} M(x) = 0 \Rightarrow \frac{\partial \phi}{\partial x} = 0.$$

Proposition 2.13. [18] *The Lie algebra rank condition is necessary and sufficient for the controllability of* (2.1) *on* $\mathbb{R}^n_*$.

Proof. Controllability of (2.1) $\Leftrightarrow$ controllability of (2.16) $\Leftrightarrow$ transitivity of $\mathbf{G}$ $\Leftrightarrow$ for each x the dimension of $\mathfrak{g}x$ is n $\Leftrightarrow$ the Lie algebra rank condition. □

Corollary 2.2. *If* $\mathfrak{g} = \{B_1, B_2\}_\mathcal{L}$ *is transitive on* $\mathbb{R}^n_*$, *then the symmetric system* $\dot{x} = (u_1B_1 + u_2B_2)x$ *is controllable on* $\mathbb{R}^n$ *using controls taking only the values* ± 1 (bang-bang controls).

Proof. The given system is controllable by Proposition 2.13. The rest is just the last paragraph of the proof of Proposition 2.7. This result recovers an early theorem of Kučera [168]. □

The Lie rank has a maximal value on $\mathbb{R}^n_*$; this value, called the generic rank of $\mathfrak{g}$, is the integer

$$\mathrm{r}(\mathfrak{g}) := \max\{\rho_\mathrm{x}(\mathfrak{g}) \mid \mathrm{x} \in \mathbb{R}^\mathrm{n}\}.$$

If $\mathrm{r}(\mathfrak{g}) = \mathrm{n}$, we say that the corresponding Lie group $\mathbf{G}$ is weakly transitive.

Remark 2.5. In a given context a property P is said to be generic in a space S if it is in that context almost always true for points in S. In this book, S will be a linear space (or a direct product of linear spaces, like $\mathbb{R}^{n\times n} \times \mathbb{R}^n$) and "$P$ is generic" will mean that P holds except at points in S that satisfy a finite set of polynomial equations.[19]

Since the n–minors of $M(x)$ are polynomials, the rank of $M(x)$ takes on its maximal value r generically in $\mathbb{R}^n$. It is easy to find the generic rank of $\mathfrak{g}$ with a symbolic calculation using the Mathematica function `Rank[Mx]` of Table 2.2.[20] If that method is not available, get lower bounds on $\mathrm{r}(\mathfrak{g})$ by

---

[18] For the case $m = 2$ with $|u(\cdot)| = 1$ see Kučera [167, 168]; for a counterexample to the claimed convexity of the reachable, set see Sussmann [257].

[19] This context is algebraic geometry; see Section C.2.3.

[20] `NullSpace` uses Gaussian elimination, over the ring $\mathbb{Q}(x_1, \dots, x_n)$ in `Rank`.

evaluating the rank of $M(x)$ at a few randomly assigned values of x (for instance, values of a random vector with uniform density on S^{n-1}).

If $\mathrm{rank}(M(x)) = n$ for x then the generic rank is $\mathrm{r}(\mathfrak{g}) = \mathrm{n}$; if $\mathrm{rank}(M(x)) < n$ at least one knows that **G** is not transitive; and the probability that $\mathrm{rank}(M(x)) < \mathrm{r}(\mathfrak{g})$ can be made extremely small by repeated trials. An example of the use of `Rank[Mx]` to obtain $\mathrm{r}(\mathfrak{g})$ is found in the function `GetIdeal` of Table 2.4. △

Definition 2.5. The rank-k locus O_k of **G** in $\mathbb{R}^n$ is the union of all the orbits of dimension k:

$$\mathrm{O}_k := \{x \in \mathbb{R}^n | \rho_x(\mathfrak{g}) = k\}. \qquad \triangle$$

Example 2.11. The diagonal group $D^+(2, \mathbb{R})$ corresponds to the symmetric system

```
 « "TreeComp";
Examples::Usage="Examples[n]: generate a random Lie \n
subalgebra of sl(n.R); global variables: n, X, A, B, LT. ”
X = Array[x, n];
Examples[N_]:= Module[{nu=N^2, k=nu}, n=N; While[k >nu-1,
{A = Table[Random[Integer, {-1, 1}], {i, n}, {j, n}];
B = Table[Random[Integer, {-1, 1}], {i, n}, {j, n}];
LT = LieTree[{A, B},n]; k = Dim[LT]; Ln = k; }];
Print["Ln= ",Ln,", ",MatrixForm[A],",",MatrixForm[B]]; LT];

GetIdeal::Usage="GetIdeal[LT] uses X, matrix list LT."
GetIdeal[LT_] := Module[{r, Ln=Length[LT]};
Mx = Transpose[LT.X]; r=Min[m, Rank[Mx]];
MinorList = Flatten[Minors[Mx, r]];
Ideal = GroebnerBasis[MinorList, X];
Print["GenRank=", r, " Ideal= "]; Ideal]

SquareFree::Usage="Input: Ideal; SQI: Squarefree ideal."
SquareFree[Ideal_]:=Module[{i, j,
lg = Length[Ideal],Tder,TPolyg,Tred},
Tder = Transpose[Table[D[Ideal, x[i]], {i, n}]];
TPolyG = Table[Apply[PolynomialGCD, Append[Tder[[j]],
Ideal[[j]]]], {j, lg}];
Tred=Simplify[Table[Ideal[[j]]/TPolyG[[j]], j,lg]];
SQI = GroebnerBasis[Tred, Y];
Print["Reduced ideal="]; SQI];

SingPoly::usage="SingPoly[Mx]; GCD of m-minors of Mx.\n"
SingPoly[M_] :=
Factor[PolynomialGCD @@ Flatten[Minors[M, n]]]
(********* Example computation for n=3 *********)
LT=Examples[3]; Mx=Transpose[LT.X]; P=SingPoly[Mx]
GetIdeal[Mx]; SquareFree[Ideal]
```

Table 2.4. Study of Lie algebras of dimension $L < \nu - 1$ (omitting $\mathfrak{gl}(n, \mathbb{R})$ and $\mathfrak{sl}(n, \mathbb{R})$).

$$\dot{x} = \begin{bmatrix} u_1 & 0 \\ 0 & u_2 \end{bmatrix} x;\ M(x) = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix};\ \det(M(x)) = x_1 x_2,\ \text{so } \mathrm{r}(\mathfrak{g}) = 2;$$

$$\rho_x(\mathfrak{d}(2,\mathbb{R})) = \begin{cases} 2, & x_1 x_2 \neq 0 \\ 1, & x_1 x_2 = 0 \text{ and } (x_1 \neq 0 \text{ or } x_2 \neq 0) \\ 0, & x_1 = 0 \text{ and } x_2 = 0 \end{cases}.$$

The four open orbits are disjoint, so $D^+(2,\mathbb{R})$ is weakly transitive on $\mathbb{R}^n$ but not transitive on $\mathbb{R}^n_*$; this symmetric system is not controllable. △

Example 2.12. (i) Using the matrices I, J of Examples (1.7)–(1.8), $\mathbf{G} \simeq \mathbb{C}_*$;

$$\mathfrak{g} = \{\alpha I + \beta J \,|\, \alpha, \beta \in \mathbb{R}\} \sim \mathbb{C};\ M(x)(\mathfrak{g}) = \{Ix,\ Jx\} = \begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix},$$
$$\det(M(x)) = x_1^2 + x_2^2 \gg 0;\ \text{so } \rho_x(\mathfrak{g}) = 2 \text{ on } \mathbb{R}^n_*.$$

(ii) The orbits of $\mathrm{SO}(n)$ on $\mathbb{R}^n_*$ are spheres S^{n-1}; $\rho_x(\mathfrak{so}(n)) = n - 1$. △

2.5 Algebraic Geometry Computations

The task of this section is to apply symbolic algebraic computations to find the orbits and the invariant polynomials of those matrix Lie groups whose generic rank on $\mathbb{R}^n_*$ is $\mathrm{r}(\mathfrak{g}) = \mathrm{n}$ or $\mathrm{r}(\mathfrak{g}) = \mathrm{n}-1$. For the elements of real algebraic geometry used here, see Appendix C and, especially for algorithms, the excellent book by Cox et al. [67]. In order to carry out *precise* calculations with matrices and polynomials in any symbolic algebra system, see Remark 2.2.

2.5.1 *Invariant Varieties*

Any list of k polynomials $p_1, \dots, p_k$ over $\mathbb{R}$ defines an affine variety (possibly empty) where they all vanish,

$$\mathbf{V}(p_1, \dots, p_k) := \{x \in \mathbb{R}^n \,|\, p_i(x) = 0,\ 1 \le i \le k\}.$$

Please read the basic definitions and facts about affine varieties in Sections C.2–C.2.1. As [67] emphasizes, many of the operations of computational algebra use algorithms that replace the original list of polynomials with a smaller list that defines the same affine variety.

Example 2.13. A familiar affine variety is a linear subspace $\mathrm{L} \subset \mathbb{R}^n$ of dimension $n-k$ defined as the intersection of the zeros of the set of independent linear functions $\{c_1^\top x, \dots, c_k^\top x\}$. There are many other sets of polynomial functions that vanish on L and nowhere else, such as $\{(c_1^\top x)^2, \dots, (c_k^\top x)^2\}$, and all these can be regarded as equivalent; the equivalence class is a polynomial ideal in $\mathbb{R}^n$. The zero-set of the single quadratic polynomial $(c_1^\top x)^2 + \cdots + (c_k^\top x)^2$ also defines L and if $k = n$, $\mathrm{L} = \{0\}$; this would not be true if we were working on $\mathbb{C}^n$.

The product $(c_1^\top x) \cdots (c_k^\top x)$ corresponds to another affine variety, the union of k hyperplanes. △

The orbits in $\mathbb{R}^n$ of a Lie group $\mathbf{G} \subset \mathrm{GL}^+(n, \mathbb{R})$ are of dimension at most r, the generic rank. In the important case $\mathrm{r}(\mathfrak{g}) = \mathrm{n}$, there are two possibilities: if $\mathbf{G}$ has Lie rank n on $\mathbb{R}^n_*$, $n \geq 2$, there is a single n-dimensional orbit; if not, there are several open orbits. Each orbit set O_k is the finite union of zero-sets of n–minors of $M(x)$, so is an affine variety itself.

The tests for transitivity and orbit classification given in Section 2.5.2 require a list

$$\{D_1, \dots, D_k\} \quad \left(k := \frac{n!}{\ell!(n-l)!}\right)$$

of the n-minors of the LARC matrix $M(x)$ defined by (2.26); each $n \times n$ determinant D_i is a homogeneous polynomial in x of degree n.

The polynomial ideal (see page 247) $\mathcal{I}_\mathfrak{g} := \langle D_1, \dots, D_k \rangle$ in the ring $\mathbb{R}[x_1, \dots, x_n]$ has other generating sets of polynomials; if such a set is linearly independent over $\mathbb{R}$, it is called a basis for the ideal.

It is convenient to have a unique basis for $\mathcal{I}_\mathfrak{g}$. To obtain this in an effective way, one requires the imposition of a total ordering $\succ$ on the monomials $x_1^{d_1} \cdots x_n^{d_n}$; a common choice for $\succ$ is the lexicographic (*lex*) order discussed in Section C.1. With respect to $\succ$ a unique reduced Gröbner basis for a polynomial ideal can be |bfa constructed by Buchberger's algorithm. This algorithm is available (and produces the same result, up to a scalar) in all modern symbolic algebra systems. In Mathematica, it is `GroebnerBasis` whose argument is a list of polynomials. It is not necessary for our purposes to go into details. For a good two-page account of Gröbner bases and Buchberger's algorithm see Sturmfels [255]; for a more complete one, see Cox et al. [67, Ch.2 and Ch.11]. There are circumstances where other total orderings speed up the algorithm.

In Table 2.4, we have a Mathematica script to find the polynomial ideal $\mathcal{I}_\mathfrak{g}$. The corresponding affine variety[21] is the singular locus $\mathrm{V}_0 = \mathbf{V}(\mathcal{I}_\mathfrak{g})$. If $\mathrm{V}_0 = \{0\}$ then $\mathfrak{g}$ is transitive on $\mathbb{R}^n_*$. The ideal $\langle x_1, \dots, x_n \rangle$ corresponds not only to $\{0\}$, but also to $\langle x^\top Q x \rangle$ for $Q \gg 0$ and to higher-degree positive definite polynomials.

The greatest common divisor (GCD) of $\mathcal{I}_\mathfrak{g} = \langle p_1, p_2, \dots, p_k \rangle$ will be useful in what follows. Since it will be used often, we give it a distinctive symbol:

$$\mathcal{P}(x) := \mathrm{GCD}(p_1, \dots, p_k)$$

[21] See Appendix C for the notation.

will be called the principal polynomial of $\mathfrak{g}$. $\mathcal{P}$ is homogeneous in x, with degree at most n. If all $D_i = 0$, then $\mathcal{P}(x) = 0$.

Remark 2.6. The GCD of polynomials in one variable can be calculated by Euclid's algorithm[22] but that method fails for $n > 1$. Cox et al. [67, pg. 187] provides an algorithm for the least common multiple (LCM) of a set of multivariable polynomials, using their (unique) factorizations into irreducible polynomials. Then

$$\mathrm{GCD}(p_1, \dots, p_k) = \frac{p_1 \cdots p_k}{\mathrm{LCM}(p_1, \dots, p_k)}.$$

Fast and powerful methods of polynomial factorization are available; see the implementation notes in Mathematica for `Factor[]`. △

The adjoint representation[23] of a matrix Lie group $\mathbf{G}$ of dimension ℓ can be given a representation by $\ell \times \ell$ matrices. If $Q \in \mathbf{G}$ and $B_i \in \mathbf{B}^\ell$, there is a matrix $C = \{c_{i,j}\} \in \mathbb{R}^{\ell \times \ell}$ depending only on Q such that

$$\mathrm{ad}_Q(B_i) = \sum_{j=1}^{\ell} c_{i,j} B_j. \tag{2.29}$$

The following two propositions show how to find relatively invariant polynomials for a $\mathbf{G}$ that is weakly transitive on $\mathbb{R}^n$.

Proposition 2.14. *If $\mathbf{G} \subset \mathrm{GL}^+(n, \mathbb{R})$ has dimension $\ell = n$ and is weakly transitive on $\mathbb{R}^n$, then there exists a real group character $\rho : \mathbf{G} \to \mathbb{R}$ for which $\mathcal{P}(Qx) = \rho(Q)\mathcal{P}(x)$ on $\mathbb{R}^n$ for all $Q \in \mathbf{G}$.*

Proof. The following argument is modified from one in Kimura [159, Cor. 2.20]. If $\mathbf{G}$ is weakly transitive on $\mathbb{R}^n$, $\mathcal{P}$ does not vanish identically. The adjoint representation of $\mathbf{G}$ on the basis matrices is given by the adjoint action (defined in (2.7)) $\mathrm{Ad}_Q(B_i) := QB_iQ^{-1}$ for all $Q \in \mathbf{G}$. Let $C(Q) = \{c_{i,j}\}$ be the matrix representation of Q with respect to the basis $\{B_1, \dots, B_n\}$ as in (2.29). Let $\xi = Q^{-1}x$, then

$$\begin{aligned} M(Q^{-1}x) =& Q^{-1}\left[QB_1Q^{-1}x \cdots\ QB_nQ^{-1}x\right] \\ =& Q^{-1}\left[B_1x \cdots\ B_nx\right] C(Q)^\top; \end{aligned} \tag{2.30}$$

$$\begin{aligned} \mathcal{P}(Q^{-1}x) =& |Q|^{-1} \det\left(\left[B_1x \cdots\ B_nx\right] C(Q)^\top\right) \\ =& |Q|^{-1}\mathcal{P}(x) \det(C(Q)). \end{aligned} \tag{2.31}$$

[22] Euclid's algorithm for the GCD of two natural numbers $p > d$ uses the sequence of remainders r_i defined by $p = dq_1 + r_1$, $d = r_1q_2 + r_2, \dots$; the last nonvanishing remainder is the GCD. This algorithm works for polynomials in *one* variable.

[23] For group representations and characters, see Section B.5.1.

Thus $\mathcal{P}$ is relatively invariant; $\rho(Q) := |Q|^{-1}\det(C(Q))$ is obviously a one-dimensional representation of $\mathbf{G}$. □

Suppose now that $\ell > n$; C is $\ell \times \ell$ and (2.30) becomes

$$M(Q^{-1}x) = Q^{-1}\left[B_1x \cdots B_\ell x\right] C(Q)^\top;$$

the rank of $M(Q^{-1}x)$ equals the rank of $M(x)$, so the rank is an integer-valued $\mathbf{G}$-invariant function.

Proposition 2.15. *Suppose* $\mathbf{G}$ *is weakly transitive on* $\mathbb{R}^n$. *Then*
(i) The ideal $\mathcal{I}_{\mathfrak{g}}$ *generated by the* $\binom{\ell}{n}$ *n-minors* D_i *of* $M(x)$ *is* $\mathbf{G}$*-invariant.*
(ii) $\mathcal{P}(x) = \mathrm{GCD}(\mathcal{I}_{\mathfrak{g}})$ *is relatively* $\mathbf{G}$*-invariant.*
(iii) The affine variety $V := \mathbf{V}(\mathcal{I}_{\mathfrak{g}})$ *is the union of orbits of* $\mathbf{G}$.

Proof. Choose any one of the $k := \binom{\ell}{n}$ submatrices of size $n \times n$, for example, $M_1(x) := \left[B_1x \cdots B_nx\right]$. For a given $Q \in \mathbf{G}$ let $C(Q) = \{c_{i,j}\}$, the $\ell \times \ell$ matrix from (2.29). Generalizing (2.30),

$$\begin{aligned} M_1(Q^{-1}x) &= Q^{-1}\left[QB_1Q^{-1}x \cdots QB_nQ^{-1}\right] \\ &= Q^{-1}\left[\textstyle\sum_{j=1}^{\ell} c_{1,j}B_jx \cdots \sum_{j=1}^{\ell} c_{n,j}B_jx\right]; \\ &= Q^{-1}\left[B_1x \cdots B_\ell x\right]\begin{bmatrix} c_{1,1} & \cdots & c_{n,1} \\ \cdots & \cdots & \cdots \\ c_{1,\ell} & \cdots & c_{n,\ell}\end{bmatrix}. \end{aligned}$$

The determinant is multilinear, so the minor $\widehat{D}_1(x) := \det(M_1(Q^{-1}x))$ is a linear combination of n-minors of $M(x)$ whose coefficients are n-minors of $C(Q)^\top$. That is, $\widehat{D}_1(x) \in \mathcal{I}_{\mathfrak{g}}$, as do all the other n-minors of $M(Q^{-1}x)$, proving (i) and (ii). To see (iii), note that the variety V is $\mathbf{G}$-invariant. □

For example, if $n = 2$, $\ell = 3$, one of the 2–minors of $M(x)$ is

$$\left|B_2Q^{-1}x\ B_3Q^{-1}x\right| =$$
$$\frac{1}{|Q|}\left(\begin{vmatrix} c_{2,1} & c_{2,2} \\ c_{3,1} & c_{3,2}\end{vmatrix}\left|B_1x\ B_2x\right| + \begin{vmatrix} c_{2,2} & c_{3,2} \\ c_{2,3} & c_{33}\end{vmatrix}\left|B_2x\ B_3x\right| + \begin{vmatrix} c_{2,1} & c_{3,1} \\ c_{2,3} & c_{3,3}\end{vmatrix}\left|B_1x\ B_3x\right|\right).$$

The results of Propositions 2.14, 2.15, and 2.11 can now be collected as follows.

(i) If $\mathbf{G}$ given by $\mathfrak{g} := \{\mathbf{a}_1, \ldots, \mathbf{a}_m\}_{\mathcal{L}}$ is weakly transitive on $\mathbb{R}^n$ then the polynomial solutions of (2.25) belong to the ideal $\mathcal{I}_{\mathfrak{g}}$ generated by the n-minors $D_i(x)$.

(ii) For any element $Q \in \mathbf{G}$ and any $x \in \mathbb{R}^n_*$, let $y = Qx$; if $\rho_x(\mathfrak{g}) = k$ then $\rho_y(\mathfrak{g}) = k$. This, at last, shows that O_k is partitioned into orbits of dimension k. Note that $\mathrm{V}_0 = \bigcup_{0 \le k < n} \mathrm{O}_k$.

(iii) The zero-sets of every solution of the set of equations

$$\mathbf{a}_1\psi(x) = 0, \ldots, \ \mathbf{a}_m\psi = 0$$

are subsets of the affine variety V defined by the polynomial invariants of $\mathfrak{g}$ of degree no more than n.

Example 2.14. From Theorem D.1 of Appendix D for each Lie algebra of Type I in its canonical representation $\mathcal{P}(x) = x_1^2 + \cdots + x_n^2$, so the singular set is $\{0\}$ as one should expect. Each of the Type I Lie algebras is of the form $\mathfrak{g} \oplus \mathfrak{c}$ where $\mathcal{P}$ is invariant for $\mathfrak{g}$ and relatively invariant for the center $\mathfrak{c}$. △

2.5.2 *Tests and Criteria*

Given a symmetric bilinear control system, this section is a guide to methods of ascertaining its orbits, using what has been established previously. Assume that the matrices $B_1, \ldots, B_m$ have entries in $\mathbb{Q}$; the algorithms described here all preserve that property for matrices and polynomials. The following tests can be applied using built-in functions and those in Tables 2.2 and 2.4.

Given (2.1), use `LieTree` to generate a basis $\{B_1, \ldots, B_m, \ldots, B_\ell\}$ for $\mathfrak{g} = \mathbf{B}^m_{\mathcal{L}}$. Construct the matrix[24]

$$M(x) := \begin{bmatrix} B_1x & B_2x & \cdots & B_mx & \cdots & B_\ell x \end{bmatrix}.$$

1. If $\ell = \nu$ then $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{R})$; if $\operatorname{tr} B_i = 0$, $i \in 1, \ldots, m$ and $\ell = \nu - 1$ then $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{R})$ by its maximality; in either case $\mathfrak{g}$ is transitive. If (n, ℓ) does not belong to any Lie algebra listed in Section D.2, $\mathfrak{g}$ is not transitive. See Table D.2 for the pairs (n, ℓ), $n \leq 9$ and $n = 16$. It must be reiterated here that for n odd, $n \neq 7$, the only transitive Lie algebras are $GL^+(n, \mathbb{R})$, $SL(n, \mathbb{R})$, and $SO(n) \times \mathfrak{c}(n, \mathbb{R})$.
2. Compute the generic rank $\mathrm{r}(\mathfrak{g})(\mathfrak{g})$ by applying the `Rank` function of Table 2.2 to $M(x)$; the maximal dimension of the orbits of $\mathbf{G}$ is $\mathrm{r}(\mathfrak{g})$, so a necessary condition for transitivity is $\mathrm{r}(\mathfrak{g}) = \mathrm{n}$.
3. Let $\nu := n^2$ as usual. From the r-minors of $M(x)$, which are homogeneous of degree r, a Gröbner basis $\langle p_1, \ldots, p_s \rangle$ is written to the screen by `GetIdeal`.[25] If $r > 3$, it may be helpful to use `SquareFree` to find an ideal `Red` of lower degree. The affine variety corresponding to `Red` is the singular variety V_0. The substitution rule (C.2) can be used for additional simplification, as our examples will hint at. If `Red` is $\langle x_1, \ldots, x_n \rangle$, it corresponds to the variety $\{0\} \subset \mathbb{R}^n$ and $\mathfrak{g}$ is transitive.
4. If $\ell \geq n$ then the GCD of $\mathcal{I}_{\mathfrak{g}}$, the singular polynomial $\mathcal{P}$, is defined. If $\mathcal{P} = 0$ then $\mathbf{G}$ cannot be transitive.

[24] Boothby and Wilson [32] pointed out the possibility of looking for common zeros of the n-minors of $M(x)$.

[25] There are $\binom{\ell}{n}$ n-minors of $M(x)$, so if ℓ is large it is worthwhile to begin by evaluating only a few of them to see if they have only 0 as a common zero.

5. In the case $r = n-1$, the method shown in Example 2.9 is applicable to look for **G**-invariant polynomials $p(x)$. For Euler's operator $\mathbf{e}$, $\mathbf{e}p(x) = rp(x)$; so if $\mathbf{a}p = 0$ then $(\mathbf{e}+\mathbf{a})p(x) = rp(x)$. For that reason, if $I \notin \mathfrak{g}$ define the augmented Lie algebra $\mathfrak{g}^+ := \mathfrak{g} \oplus \mathbb{R}I$, the augmented matrix $M(x)^+ = \begin{bmatrix} x\ B_1x\ B_2x\ \ldots\ B_\ell x \end{bmatrix}$ and $\mathcal{P}^+$, the GCD of the n-minors of $M(x)^+$. If $\mathcal{P}^+$ does not vanish, its level sets are orbits of **G**.
6. If $\deg(\mathcal{P}) > 0$, the affine variety $\mathrm{V}_0 := \{x \in \mathbb{R}^n | \mathcal{P}(x) = 0\}$ has codimension of at least one; it is nonempty since it contains $\{0\}$.
7. If the homogeneous polynomial $\mathcal{P}$ is positive definite, $\mathrm{V}_0 = \{0\}$, and transitivity can be concluded. In real algebraic geometry, replacements like $\{p_1^2 + p_2^2\} \to \{p_1, p_2, p_1^2 + p_2^2\}$ can be used to simplify the computations.
8. If $\mathcal{P}(x) = 1$ then $\dim \mathrm{V}_0 < n-1$, which means that $\mathrm{V}_0^c := \mathbb{R}^n \setminus \mathrm{V}_0$ is connected and there is a single open orbit, so the system is controllable on V_0^c. The Gröbner basis $\langle p_1, \ldots, p_k \rangle$ must be examined to find V_0. If $\mathrm{V}_0 = \{0\}$ then **G** is transitive on $\mathbb{R}^n_*$.
9. Evaluating the GCD of the $k := \binom{\ell}{n}$ minors of $M(x)$ with large ℓ is time-consuming, and a better algorithm would evaluate it in a sequential way after listing the minors:
   $q_1 := \mathrm{GCD}(D_1, D_2)$, $q_2 := \mathrm{GCD}(q_1, D_3), \ldots$;
   if $\mathrm{GCD}(q_i, D_{i+1}) = 1$ then stop, because $\mathcal{P} = 1$;
   else the final q_k has degree larger than 0.
10. The task of finding the singular variety $\mathcal{V}_0$ is sometimes made easier by finding a square-free ideal that will define it; that can be done in Mathematica by `GetSqFree` of Table 2.4. (However, the canonical ideal that should be used to represent V_0 is the real radical $\sqrt[R]{\mathcal{I}_\mathfrak{g}}$ discussed in Section C.2.2.) Calculating radicals is most efficiently accomplished with Singular [110]. △

Mathematica's built-in function `PolynomialGCD` takes the greatest common divisor (GCD) of any list of multivariable polynomials. For those using Mathematica, the operator `MinorGCD` given in Table 2.4 can be used to find $\mathcal{P}(x)$ for a given bilinear system.

The computation of $\mathcal{P}$ or $\mathcal{P}^+$, the ideal $\mathcal{I}_\mathfrak{g}$, and the square-free ideal $\mathcal{I}'_\mathfrak{g}$ are shown in Table 2.4 with a Mathematicascript. The square-free ideal `Rad` corresponds to the singular variety V_0. To check that `Rad` is actually the radical ideal $\sqrt{\mathbf{I}}$, which is canonical, requires a real algebraic geometry algorithm that seems not to be available.

For $n = 2, 3, 4$, the computation of $\mathcal{P}$ and the Gröbner basis takes only a few minutes for randomly generated matrices and for the transitive Lie algebras of Appendix D.

Example 2.15. Let us find the orbits for $\dot{x} = uAx + vBx$ with

$$Ax = \begin{bmatrix} x_3 \\ 0 \\ x_1 \end{bmatrix}, \ Bx = \begin{bmatrix} x_2 \\ -x_1 \\ 0 \end{bmatrix}. \quad [A, B]x = \begin{bmatrix} 0 \\ x_3 \\ x_2 \end{bmatrix}; \ \mathcal{P}(x) = \begin{vmatrix} x_3 & x_2 & 0 \\ 0 & -x_1 & x_3 \\ x_1 & 0 & x_2 \end{vmatrix} \equiv 0.$$

Its Lie algebra $\mathfrak{g} = \mathfrak{so}(1, 2)$ has dimension three. Its augmented Lie algebra (see page 64) $\mathfrak{g}^+ = \{A, B, I\}_{\mathcal{L}}$ is four-dimensional and we have

$$M^+(x) = \begin{bmatrix} x_3 & x_2 & 0 & x_1 \\ 0 & -x_1 & x_3 & x_2 \\ x_1 & 0 & x_2 & x_3 \end{bmatrix}; \ \mathcal{P}^+(x) = x_3^2 - x_2^2 - x_1^2;$$

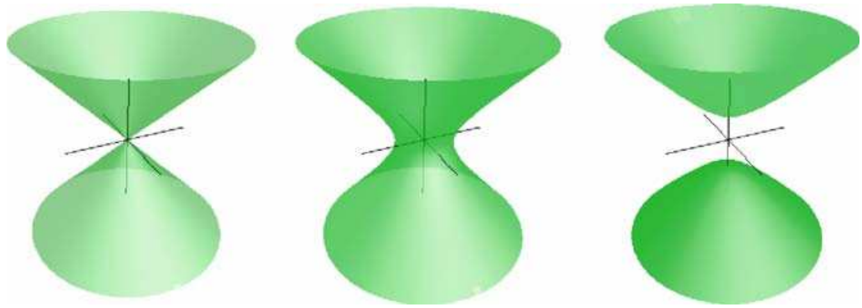
the varieties $\mathrm{V}_c := \{x|\, x_3^2 - x_2^2 - x_1^2 = c, \ c \in \mathbb{R}\}$ contain orbits of $\mathbf{G} = \mathrm{SO}(2, 1)$. Figure 2.2(a) shows three distinct orbits (one of them is $\{0\}$); 2.2(b) shows one and 2.2(c) shows two. △

Example 2.16 will illustrate an important point: saying that the closure of the set of orbits of $\mathbf{G}$ is $\mathbb{R}^n$ is not the same as saying that $\mathbf{G}$ has an open orbit whose closure is $\mathbb{R}^n$ (weak transitivity). It also illustrates the fact that the techniques discussed here can be applied to control systems whose Lie groups are not closed.

Example 2.16. Choose an integer $k > 0$ and let

$$B_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -k & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

If $\sqrt{k}$ is not an integer, then as in Example 2.5 the closure of the one-parameter group $\mathbf{G} = \exp(\mathbb{R}B_1)$ is a torus subgroup $T^2 \subset \mathrm{GL}^+(4, \mathbb{R})$. Each orbit of $\mathbf{G}$ (given by $\dot{x} = B_1 x, \ x(0) \in \mathbb{R}^4_*$) is dense in a torus in $\mathbb{R}^4$, $\{x|\, p_1(x) = p_1(\xi), \, p_2(x) = p_2(\xi)\}$, where $p_1(x) := x_1^2 + x_2^2$, $p_2(x) := kx_3^2 + x_4^2$. Now construct $M(x)^\top$, the LARC matrix transposed, for the three-input symmetric control system



(a) (b) (c)

Fig. 2.2. Orbits of SO(2, 1) for Example 2.15.

$$\dot{x} = (u_1 B_1 + u_2 B_2 + u_3 B_3)x,\ x(0) = \xi; \tag{2.32}$$

$$M(x)^\top = \begin{bmatrix} x_2 & -x_1 & x_4 & -2x_3 \\ x_1 & x_2 & 0 & 0 \\ 0 & 0 & x_3 & x_4 \end{bmatrix}.$$

The ideal generated by the 3×3 minors of $M(x)$ has a Gröbner basis $\mathcal{I}_{\mathfrak{g}} := \langle -x_2 p_2(x), -x_1 p_2(x), x_4 p_1(x), x_3 p_1(x) \rangle$. The corresponding affine variety is $\mathrm{V}(\mathcal{I}_{\mathfrak{g}}) = \mathcal{V}\langle p_1 \rangle \bigcup \mathcal{V}\langle p_2 \rangle$, the union of two 2-planes.

Let $\mathbf{G}$ denote the matrix Lie group generated by (2.32), which is three dimensional and Abelian. The generic rank of $M(x)$ is $\mathrm{r}(\mathfrak{g}) = 3$, so a three-dimensional orbit $\mathbf{G}\xi$ of (2.32) passes through each ξ that is not in P_1 or P_2. If $\sqrt{k}$ is an integer, $\mathbf{G} = \mathbb{R}_+ \times \mathbb{R}_+ \times T^1$. If $\sqrt{k}$ is not an integer, then the closure of $\mathbf{G}$ in $\mathrm{GL}(4, \mathbb{R})$ is (see (2.14) for $\boldsymbol{\alpha}$)

$$\mathbb{R}_+ \times \mathbb{R}_+ \times T^2 \simeq \boldsymbol{\alpha}(\mathbb{C}_*) \times \boldsymbol{\alpha}(\mathbb{C}_*);$$

it has dimension four and is weakly transitive. △

Example 2.17. The symmetric matrix control systems

$$\dot{X} = u_1 B_1 X + u_2 B_2 X \text{ on } \mathbb{R}^{n \times n} \tag{2.33}$$

have a Kronecker product representation on $\mathbb{R}^\nu$ ($\nu := n^2$; see Section A.3.2 and (A.12)). For the case $n = 2$, write $z = \mathrm{col}(z_1, z_3, z_2, z_4)$,

$$X := \begin{bmatrix} z_1 & z_3 \\ z_2 & z_4 \end{bmatrix}, \text{ that is, } z = X^\flat; \quad \dot{z} = u_1(I \otimes B_1)z + u_2(I \otimes B_2)z. \tag{2.34}$$

Suppose, for example, that the group $\mathbf{G}$ of transition matrices for (2.34) is $\mathrm{GL}(2, \mathbb{R})$, whose action is transitive on $\mathrm{GL}^+(2, \mathbb{R})$. Let $q(z) := \det(X) = z_1 z_4 - z_2 z_3$. The lifted system (2.34) is transitive on the open subset of $\mathbb{R}^4$ defined by $\{z | q(z) > 0\}$, and $\{z | q(z) = 0\}$ is the single invariant variety. The polynomial $\mathcal{P}(z) := q(z)^2$ for (2.34).

Many invariant varieties are unions of quadric hypersurfaces, so it is interesting to find counterexamples. One such is the instance $n = 3$ of (2.33) when the transition group is $\mathrm{GL}(3, \mathbb{R})$. Its representation by $z = X^\flat$, as a bilinear system on $\mathbb{R}^9$, is an action of $\mathrm{GL}(3, \mathbb{R})$ on $\mathbb{R}^9$ for which the polynomial $\det(X)$ is an invariant cubic in z with no real factors. △

Exercise 2.3. Choose some nonlinear functions h and vector fields f, g and experiment with the following script.

```
 n=Input["n=?"]; X=Array[x,n];
LieDerivative::usage="LieDerivative[f,h,X]: Lie derivative
of h[x] for vector field f(x).";
LieDerivative[f_,h_,X_]:=Sum[f[[i]]*D[h,X[[i]]],{i,1,n}];
```

```
LieBracket::usage="LieBracket[f,g,X]: Lie bracket [f,g]\n
at x of the vector fields f and g.";
LieBracket[f_,g_,X_]:=
Sum[D[f,X[[j]]]*g[[j]]-D[g,X[[j]]]*f[[j]],{j,1,n}]
```

2.6 Low-Dimensional Examples

Vector field (rather than matrix) representations of Jacobson's list [143, I.4] of the Lie algebras whose dimension ℓ is two or three are given in Sections 2.6.1 and 2.6.2. The Lie bracket of two vector fields is (B.6); it can be calculated with the Mathematica function `LieBracket` in Exercise 2.3.

These Lie algebras are defined by their bases, respectively $\{a, b\}$ or $\{a, b, c\}$, the brackets that relate them, and the axioms Jac.1–Jac.3. The lists in Sections 2.6.1 and 2.6.2 give the relations; the basis of a faithful vector field representation; the generic rank $\mathrm{r}(\mathfrak{g})$; and if $\mathrm{r}(\mathfrak{g}) = \mathrm{n}$, the singular polynomial $\mathcal{P}$. If $\mathrm{r}(\mathfrak{g}) = \mathrm{n} - 1$ the polynomial $\mathcal{P}^+$ is given; it is the GCD of the $n \times n$ minors of $M(x)^+$, is relatively invariant, and corresponds to an invariant variety of dimension $n - 1$.

2.6.1 *Two-Dimensional Lie Algebras*

The two-dimensional Abelian Lie algebra $\mathfrak{a}_2$ has one defining relation, $[a, b] = 0$, so Jac.1–Jac.3 are satisfied. A representation of $\mathfrak{a}_2$ on $\mathbb{R}^2$ is of the form $\{\mathbb{R}I \oplus \mathbb{R}A\}$ where $A \in \mathbb{R}^{2\times 2}$ is *arbitrary*, so examples of $\mathfrak{a}_2$ include

$$\mathfrak{d}_2 : a = \begin{bmatrix} x_1 \\ 0 \end{bmatrix},\ b = \begin{bmatrix} 0 \\ x_2 \end{bmatrix},\ \mathcal{P}(x) = x_1 x_2;$$
$$\boldsymbol{\alpha}(\mathbb{C}) : a = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},\ b = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix},\ \mathcal{P}(x) = x_1^2 + x_2^2;$$
$$\text{and } \mathfrak{so}(1,1) : a = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},\ b = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix};\ \mathcal{P}(x) = x_1^2 - x_2^2.$$

The two-dimensional non-Abelian Lie algebra $\mathfrak{n}_2$ has one relation $[e, f] = \lambda e$ (for any $\lambda \neq 0$); it can, by the linear mapping $a = e$, $b = f/\lambda$, be put into a standard form with relation $[a, b] = a$. This standard $\mathfrak{n}_2$ can be represented on $\mathbb{R}^2$ by any member of the one-parameter family $\{\mathfrak{n}_2(\alpha),\ \alpha \in \mathbb{R}\}$ of isomorphic (but *not* conjugate) Lie algebras:

$$\mathfrak{n}_2(\alpha) :\ a = \begin{bmatrix} x_2 \\ 0 \end{bmatrix},\ b_s = \begin{bmatrix} (\alpha - 1)x_1 \\ s x_2 \end{bmatrix};\ \mathcal{P}(x) = \alpha x_2^2.$$

Each representation on $\mathbb{R}^2$ of its Lie group $N(2,\mathbb{R})$ with $\alpha \neq 0$ is weakly transitive. If $s = 0$, its orbits in $\mathbb{R}^2_*$ are the lines with x_2 constant. If $s \neq 0$, they are

$$\begin{bmatrix}\mathbb{R}\\ \mathbb{R}_+\end{bmatrix}, \begin{bmatrix}\mathbb{R}\\ \mathbb{R}_-\end{bmatrix}, \begin{bmatrix}\mathbb{R}_+\\ 0\end{bmatrix}, \begin{bmatrix}\mathbb{R}_-\\ 0\end{bmatrix}.$$

2.6.2 Three-Dimensional Lie Algebras

Jacobson [143, I.4] (IIIa–IIIe) is the standard list of relations for each conjugacy class of real three-dimensional Lie algebras. From this list, the vector field representations below have been calculated as an example. The structure matrices A in class 4 are the real canonical matrices for $GL(2,\mathbb{R})/\mathbb{R}_*$. The representations here are three dimensional except for $\mathfrak{g}_7 = \mathfrak{sl}(2,\mathbb{R}) \subset \mathfrak{gl}(2,\mathbb{R})$.

$$(\mathfrak{g}_1)\ [a,b] = 0,\ [b,c] = 0,\ [a,c] = 0;\ a = \begin{bmatrix}x_1\\0\\0\end{bmatrix}, b = \begin{bmatrix}0\\x_2\\0\end{bmatrix},\ c = \begin{bmatrix}0\\0\\x_3\end{bmatrix};$$

$$r = 3;\ \mathcal{P}(x) = x_1x_2x_3.$$

$$(\mathfrak{g}_2)\ [a,b] = 0,\ [b,c] = a,\ [a,c] = 0;\quad a = \begin{bmatrix}0\\0\\x_1\end{bmatrix},\ b = \begin{bmatrix}0\\x_1\\0\end{bmatrix},\ c = \begin{bmatrix}0\\0\\x_2\end{bmatrix};$$

$$\mathfrak{g}_2 = \mathfrak{aff}(2,\mathbb{R}),\ r = 2;\quad \mathcal{P}^+(x) = x_1.$$

$(\mathfrak{g}_3)$ $[a,b] = a,\ [a,c] = 0,\ [b,c] = 0;\quad \mathfrak{g}_3 = \mathfrak{n}_2 + \mathbb{R}c;$ See Fig. 2.3.

$$a = \begin{bmatrix}0\\-2x_3\\x_1\end{bmatrix},\ b = \begin{bmatrix}0\\2x_2\\x_3\end{bmatrix},\ c = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix};\ r = 3;\ \mathcal{P}(x) = x_1(x_1x_2 + x_3^2).$$

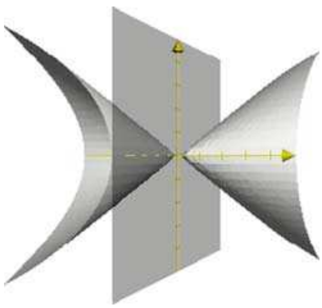


Fig. 2.3. The singular variety V_0 for $\mathfrak{g}_3$: union of a cone and a plane tangent to it.

($\mathfrak{g}_4$) $[a,b] = 0$, $[a,c] = \alpha e + \beta b$, $[b,c] = \gamma e + \delta f$; $r = 2$; three canonical representations of $\mathfrak{g}_4$ correspond to canonical forms of $P := \begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix}$:

($\mathfrak{g}_{4a}$) $P = \begin{bmatrix}1 & 0\\ 0 & \delta\end{bmatrix}$, $\delta \neq 0; a = \begin{bmatrix}0\\0\\x_1\end{bmatrix} b = \begin{bmatrix}0\\0\\x_2\end{bmatrix} c = \begin{bmatrix}x_1\\ \delta x_2\\ 0\end{bmatrix}$; $\mathcal{P}^+(x) = x_1 x_2$.

($\mathfrak{g}_{4b}$) $P = \begin{bmatrix}1 & \beta\\ 0 & 1\end{bmatrix}$, $\beta \neq 0$; $a = \begin{bmatrix}0\\0\\x_1\end{bmatrix}$, $b = \begin{bmatrix}0\\0\\x_2\end{bmatrix}$, $c = \begin{bmatrix}x_1 + \beta x_2\\ x_2\\ 0\end{bmatrix}$; $\mathcal{P}^+(x) = x_2$.

($\mathfrak{g}_{4c}$) $P = \begin{bmatrix}\alpha & \beta\\ -\beta & \alpha\end{bmatrix}$, $\alpha^2 + \beta^2 = 1$; $a = \begin{bmatrix}0\\0\\x_2\end{bmatrix}$, $b = \begin{bmatrix}0\\0\\-x_1\end{bmatrix}$, $c = \begin{bmatrix}\alpha x_1 + \beta x_2\\ -\beta x_1 + \alpha x_2\\ 0\end{bmatrix}$;

$\mathcal{P}^+(x) = x_1^2 + x_2^2$; $\mathcal{V}_0$ is the x_3-axis.

There are three (Section B.2) simple Lie algebras of dimension three:

($\mathfrak{g}_5$) $[a,b] = c$, $[b,c] = a$, $[c,a] = b$; $a = \begin{bmatrix}x_2\\ -x_1\\ 0\end{bmatrix}, b = \begin{bmatrix}x_3\\ 0\\ -x_1\end{bmatrix}$, $c = \begin{bmatrix}0\\ -x_3\\ x_2\end{bmatrix}$;

$\mathfrak{g}_5 = \mathfrak{so}(3)$, $r = 2$; $\mathcal{P}^+(x) = x_1^2 + x_2^2 + x_3^2$.

($\mathfrak{g}_6$) $[a,b] = c$, $[b,c] = -a$, $[c,a] = b$; $a = \begin{bmatrix}x_2\\ -x_1\\ 0\end{bmatrix}, b = \begin{bmatrix}x_3\\ 0\\ x_1\end{bmatrix}$, $c = \begin{bmatrix}0\\ x_3\\ x_2\end{bmatrix}$;

$\mathfrak{g}_6 = \mathfrak{so}(2,1)$, $r = 2$; $\mathcal{P}^+(x) = x_1^2 + x_2^2 - x_3^2$.

($\mathfrak{g}_7$) $[a,b] = c$, $[b,c] = 2b$, $[c,a] = 2a$; $a = \begin{bmatrix}x_2\\ 0\end{bmatrix}$, $b = \begin{bmatrix}0\\ x_1\end{bmatrix}$, $c = \begin{bmatrix}x_1\\ -x_2\end{bmatrix}$;

$\mathfrak{g}_7 = \mathfrak{sl}(2,\mathbb{R})$; $r = 2$, $\mathcal{P}(x) = 1$.

The representations of $\mathfrak{g}_7$ on $\mathbb{R}^3$ (IIIe in [143, I.4]) are left as an exercise.

2.7 Groups and Coset Spaces

We begin with a question. What sort of manifold can be acted on transitively by a given Lie group **G**?

Reconsider matrix bilinear systems (2.2): each of the B_i corresponds to a right-invariant vector field $\dot{X} = B_i X$ on $\mathrm{GL}^+(n,\mathbb{R}) \subset \mathbb{R}^{n\times n}$. It is also right-invariant on **G**. Thus, system (2.2) is controllable (transitive) acting on its own group **G**.

Suppose we are now given a nontrivial normal subgroup **H** $\subset$ **G** (i.e., **HG** = **GH**); then the quotient space **G**/**H** is a Lie group (see Section B.5.3) which is acted on transitively by **G**. Normality can be tested at the Lie algebra level: **H** is a normal subgroup if and only if its Lie algebra $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. An example is SO(3)/SO(2) = T^1, the circle group; it is easy to see that $\mathfrak{so}(2)$ is an ideal in $\mathfrak{so}(3)$.

If a closed subgroup (such as a finite group) $\mathbf{H} \subset \mathbf{G}$ is not normal, then (see Section B.5.7) the quotient space $\mathbf{G}/\mathbf{H}$, although not a group, can be given a Hausdorff topology compatible with the action of **G** and is called a homogeneous space or coset space. A familiar example is to use a subgroup of translations $\mathbb{R}^k \subset \mathbb{R}^n$ to partition Euclidean space: $\mathbb{R}^{n-k} \simeq \mathbb{R}^n/\mathbb{R}^k$.

For more about homogeneous spaces, see Section 6.4 and Helgason [120]; they were introduced to control theory in Brockett [36] and Jurdjevic and Sussmann [151]. The book of Jurdjevic [147, Ch. 6] is a good source for the control theory of such systems with examples from classical mechanics. Some examples are given in Chapter 6.

The examples of homogeneous spaces of primary interest in this chapter are of the form $\mathbb{R}^n_* \simeq \mathbf{G}/\mathbf{H}$ where **G** is one of the transitive Lie groups. For a given matrix Lie group **G** transitive on $\mathbb{R}^n_*$ and any $\xi \in \mathbb{R}^n_*$, the set of matrices $\mathbf{H}_\xi := \{Q \in \mathbf{G} |\, Q\xi = \xi\}$ is (i) a subgroup of **G** and (ii) closed, being the inverse image of ξ with respect to the action of **G** on $\mathbb{R}^n$, so $\mathbf{H}_\xi$ is a Lie subgroup, the isotropy subgroup of **G** at ξ.[26]

An isotropy subgroup has, of course, a corresponding matrix Lie algebra $\mathfrak{h}_\xi := \{A \in \mathfrak{g} |\, A\xi = 0\}$ called the isotropy subalgebra of $\mathfrak{g}$ at ξ, unique up to conjugacy: if $T^{-1}\zeta = \xi$ then $\mathfrak{h}_\xi = T\mathfrak{h}_\zeta T^{-1}$.

Example 2.18. Some isotropy algebras for transitive Lie algebras are given in Appendix D. For $\xi = \delta_k$ (see (A.1)), the isotropy subalgebra of a transitive matrix Lie algebra $\mathfrak{g}$ is the set of matrices in $\mathfrak{g}$ whose kth column is 0_n. Thus for $\mathfrak{g} = \mathfrak{sl}(3, \mathbb{R})$ and $\xi = \delta_3$

$$\mathfrak{h}_\xi = \left\{ \begin{bmatrix} u_1 & u_2 & 0 \\ u_3 & -u_1 & 0 \\ u_4 & u_5 & 0 \end{bmatrix} \middle|\, u_1, \ldots, u_5 \in \mathbb{R} \right\}. \qquad \triangle$$

Example 2.19 (Rotation Group, I). Other homogeneous spaces, such as spheres, are encountered in bilinear system theory. The rotation group SO(3) is a simple group (see Section B.2.2 for *simple*). A usual basis of its matrix Lie algebra $\mathfrak{so}(3)$ is

$$B_1 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; B_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \; B_3 = [B_1, B_2] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

$$\text{Then } \mathfrak{so}(3) \simeq \{B_1, B_2\}_{\mathcal{L}} = \text{span}(B_1, B_2, B_3); \quad \dot{X} = \sum_{i=1}^{3} u_i B_i X$$

[26] Varadarajan [282] uses *stability subgroup* instead of isotropy subgroup.

is a matrix control system on $\mathbf{G} = \mathrm{SO}(3)$. The isotropy subgroup at δ_1 is SO(2), represented as the rotations around the x_1 axis; the corresponding homogeneous space is the sphere $S^2 \simeq \mathrm{SO}(3)/\mathrm{SO}(1)$. The above control system is used in physics and engineering to describe controlled motions like aiming a telescope. It also describes the rotations of a coordinate frame (rigid body) in $\mathbb{R}^3$, since an oriented frame can be represented by a matrix in SO(3). △

Problem 2.1. The Lie groups transitive on spheres S^n are listed in (D.2). Some are simple groups like SO(3) (Example 2.19). Others are not; for $n > 3$ some of these non-simple groups must have Lie subgroups that are *weakly transitive* on S^n. Their orbits on S^n might, as an exercise in real algebraic geometry, be studied using #4 in Section 2.5.2. △

2.8 Canonical Coordinates

Coordinate charts for Lie groups are useful in control theory and in physics. Coordinate charts usually are local, as in the discussion of Proposition 2.6, although the group property implies that all coordinate charts are translations of a chart containing the identity. A global chart of a manifold $\mathcal{M}^n$ is a single chart that covers all of the manifold; the coordinate mapping $x : \mathcal{M}^n \to \mathbb{R}^n$ is a diffeomorphism. That is impossible if $\mathbf{G}$ is compact or is not simply connected; for instance, the torus group

$$T^n \simeq \mathrm{diag}\Big(e^{x_1 J}, \dots, e^{x_n J}\Big) \subset \mathrm{GL}(2n, \mathbb{R})$$

has local, but not global, coordinate charts $T^n \to x$. See Example 2.4.

Note: to show that a coordinate chart $x : \mathbf{G} \to \mathbb{R}^n$ on the underlying manifold of a Lie group $\mathbf{G}$ is global *for the group,* it must also be shown that on the chart the group multiplication and group inverse are C^ω.

2.8.1 *Coordinates of the First Kind*

It is easy to see that canonical coordinates of the first kind (CCK1)

$$e^{t_1 B_1 + \cdots + t_\ell B_\ell} \to (t_1, \dots, t_\ell)$$

are global if and only if the exponential map $\mathfrak{G} \to \mathbf{G}$ is one-to-one and onto. For instance, the mapping of diagonal elements of $X \in D^+(n, \mathbb{R})$ to their logarithms maps that group one-to-one onto $\mathbb{R}^n$. Commutativity of $\mathbf{G}$ is not enough for the existence of global CCK1; an example is the torus group $T^k \subset \mathrm{GL}^+(2k, \mathbb{R})$.

However, those nilpotent matrix groups that are conjugate to $\mathrm{N}(m,\mathbb{R})$ have CCK1. If $Q \in \mathrm{N}(n,\mathbb{R})$ then $Q - I$ is (in every coordinate system) a nilpotent matrix;

$$\log(Q) = (Q-I) - \frac{1}{2}(Q-I)^2 + \cdots - (-1)^{n-1}\frac{1}{n-1}(Q-I)^{n-1} \in \mathfrak{n}(n,\mathbb{R})$$

identically. Thus the mapping $X = \log(Q)$ provides a CCK1 chart, polynomial in $\{q_{i,j}\}$ and global. For most other groups, the exponential map is not onto, as the following example and Exercises 2.4 and 2.5 demonstrate.

Example 2.20 (Rotation group, II). Local CCK1 for SO(3) are given by Euler's theorem: for any $Q \in \mathrm{SO}(3)$, there is a real solution of $Q\xi = \xi$ with $\|\xi\| \le \pi$ such that $Q = \exp(\xi_1 B_3 + \xi_2 B_2 - \xi_3 B_1)$. △

Exercise 2.4. Show that the image of $\exp : \mathfrak{sl}(2,\mathbb{R}) \to \mathrm{SL}(2,\mathbb{R})$ consists of precisely those matrices $A \in \mathrm{SL}(2,\mathbb{R})$ such that $\operatorname{tr} A > 2$, together with the matrix I (which has trace 2). △

Exercise 2.5 (Varadarajan [282]). Let $\mathbf{G} = \mathrm{GL}(2,\mathbb{R})$; the corresponding Lie algebra is $\mathfrak{g} = \mathfrak{gl}(2,\mathbb{R})$. Let $X = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \in \mathbf{G}$; show $X \notin \exp(\mathfrak{g})$. △

Remark 2.7. Markus [201] showed that if matrix $A \in \mathbf{G}$, an algebraic group $\mathbf{G}$ with Lie algebra $\mathfrak{g}$, then there exists an integer N and $t \in \mathbb{R}$ and $B \in \mathfrak{g}$ such that $A^N = \exp(tB)$. The minimal power N needed here is bounded in terms of the eigenvalues of A. △

2.8.2 *The Second Kind*

Suppose that $\mathfrak{g} := \{B_1, \ldots, B_m\}_{\mathcal{L}}$ has basis $\{B_1, \ldots, B_m, B_{m+1}, \ldots, B_\ell\}$ and is the Lie algebra of a matrix Lie group $\mathbf{G}$. If $X \in \mathbf{G}$ is near the identity, its canonical coordinates of the second kind (CCK2) are given by

$$e^{t_\ell B_\ell} \cdots e^{t_2 B_2} e^{t_1 B_1} \to (t_1, t_2, \ldots, t_\ell) \tag{2.35}$$

The question then arises, when are the CCK2 coordinates $t_1, \ldots, t_\ell$ global? The answer is known in some cases.

We have already seen that the diagonal group $D^+(n,\mathbb{R}) \subset \mathrm{GL}^+(n,\mathbb{R})$ has CCK1; to get the second kind, trivially one can use factors $\exp(t_k J_k)$ where $J_k := \delta_{i,j}\delta_{i,k}$ to get

$$\exp(t_1 J_1) \cdots \exp(t_n J_n) \mapsto (t_1, t_2, \ldots, t_n) \in \mathbb{R}^n.$$

A solvable matrix Lie group $\mathbf{G}$ is a subgroup of $\mathrm{GL}^+(n,\mathbb{R})$ whose Lie algebra $\mathfrak{g}$ is solvable; see Section B.2.1. Some familiar matrix Lie groups are solvable: the positive-diagonal group $D^+(n, \mathbb{R})$, tori $T^n \simeq \mathbb{R}^n/\mathbb{Z}^n$, upper triangular groups $\mathrm{T}^u(n, \mathbb{R})$, and the unipotent groups $\mathrm{N}(n, \mathbb{R})$ (whose Lie algebras $\mathfrak{n}(n, \mathbb{R})$ are nilpotent).

A solvable Lie group $\mathbf{G}$ may contain as a normal subgroup a torus group T^k, in which case it cannot have global coordinates of any kind. Even if solvable $\mathbf{G}$ is simply connected, Hochschild [129, p. 140] has an example for which the exponential map $\mathfrak{g} \to \mathbf{G}$ is not surjective, defeating any possibility of global first-kind coordinates. However, there exist global CCK2 for simply connected solvable matrix Lie groups; see Corollary B.1; they correspond to an ordered basis of $\mathfrak{g}$ that is *compatible* with its derived series $\mathfrak{g} \supset \mathfrak{g}' \cdots \supset \mathfrak{g}'' \supset \cdots \supset \{0\}$ (see Section B.2.1) in the sense that $\mathrm{span}\{B_1\} = \mathfrak{g}'$, $\mathrm{span}\{B_1, B_2\} = \mathfrak{g}''$, and so forth as in the following example.

Example 2.21. The upper triangular group $\mathfrak{T}^u(n, \mathbb{R}) \subset \mathrm{GL}^+(n, \mathbb{R})$ has Lie algebra $\mathfrak{t}^u(n, \mathbb{R})$. There exists an ordered basis $\{B_1, \dots, B_\ell\}$ of $\mathfrak{t}^u(n, \mathbb{R})$ such that (2.35) provides a global coordinate system for $\mathfrak{T}(n, \mathbb{R})$. The mapping $\mathfrak{T}^u(n, \mathbb{R}) \to \mathbb{R}^k$ (where $k = n(n+1)/2$) can be best seen by an example for $n = 3$, $k = 6$ in which the matrices $B_1, \dots, B_6$ are given by

$$B_i := \begin{bmatrix} \delta_{i,4} & \delta_{i,2} & \delta_{i,1} \\ 0 & \delta_{i,5} & \delta_{i,3} \\ 0 & 0 & \delta_{i,6} \end{bmatrix}; \text{ then } e^{t_6 B_6} \cdots e^{t_2 B_2} e^{t_1 B_1} = \begin{bmatrix} e^{t_4} & t_2 e^{t_4} & t_1 e^{t_4} \\ 0 & e^{t_5} & t_3 e^{t_5} \\ 0 & 0 & e^{t_6} \end{bmatrix}.$$

The mapping $\mathfrak{T}^u(3, \mathbb{R}) \to \mathbb{R}^6$ is a C^ω diffeomorphism. △

Example 2.22 (Rotation group, III). Given $Q \in \mathrm{SO}(3)$, a local set of CCK2 is given by $Q = \exp(\phi_1 B_1)\exp(\phi_2 B_2)\exp(\phi_3 B_3)$; the coordinates ϕ_i are called Euler angles[27] and are valid in the regions $-\pi \le \phi_1 < \pi$, $-\pi/2 < \phi_2 < \pi/2$, $-\pi \le \phi_3 < \pi$. The singularities at $|\phi_2| = \pi/2$ are familiar in mechanics as "gimbal lock" for gyroscopes mounted in three-axis gimbals. △

2.9 Constructing Transition Matrices

If the Lie algebra $\mathfrak{g}$ generated by system

$$\dot{X} = \sum_{i=1}^{m} u_i B_i X, \; X(0) = I \text{ has basis } \{B_1, \dots, B_\ell\}, \text{ set} \tag{2.36}$$

$$X(t; u) = \mathrm{e}^{\phi_1(t) B_1} \dots \mathrm{e}^{\phi_\ell(t) B_\ell}. \tag{2.37}$$

---

[27] The ordering of the Euler angle rotations varies in the literature of physics and engineering.

If $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{R})$ then for a short time and bounded $\|u(\cdot)\|$ any matrix trajectory in **G** starting at I can be represented by (2.37), as if the ϕ_i are second canonical coordinates for **G**.[28]

Proposition 2.16 (Wei and Norman [285]). *For small t, the solutions of system* (2.36) *can be written in the form* (2.37). *The functions $\phi_1, \dots, \phi_\ell$ satisfy*

$$\Psi(\phi_1, \dots, \phi_\ell) \begin{bmatrix} \dot{\phi}_1 \\ \vdots \\ \dot{\phi}_\ell \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ u_\ell \end{bmatrix}, \tag{2.38}$$

where $u_i(t) = 0$ for $m < i \le \ell$; the matrix Ψ is C^ω in the ϕ_i and in the structure constants $\gamma^i_{j,k}$ of $\mathfrak{g}$. △

Examples 2.21–2.26 and Exercise 2.6 show the technique, which uses Proposition 2.2 and Example 2.3. For a computational treatment, expressing $e^{t\,\mathrm{ad}_A}$ as a polynomial in ad_A, see Altafini [6].[29]

To provide a relatively simple way to calculate the ϕ_i, a product representation (2.37) with respect to Philip Hall bases on m generators $\mathcal{PH}(m)$ (see page 227 for formal Philip Hall bases) can be obtained as a bilinear specialization of Sussmann [266]. If the system Lie algebra is nilpotent, the Hall-Sussmann product (like the bilinear Example 2.24) gives a global representation with a finite number of factors; it is used for trajectory construction (motion planning) in Lafferriere and Sussmann [172], and in Margaliot [199] to obtain conditions for controllability when Ω is a finite set (bang-bang control). The relative simplicity of product representations and CCK2 for a nilpotent Lie algebra $\mathfrak{g}$ comes from the fact that there exists k such that for any $X, Y \in \mathfrak{g}$, $\mathrm{ad}^k_X(Y) = 0$. Then $\exp(tX)Y\exp(-tX)$ is a polynomial in t of degree $k-1$.

Example 2.23 (Wei and Norman [284]). Let $\dot{X} = (uA + vB)X$, $X(0) = I$ with $[A, B] = B$. The Lie algebra $\mathfrak{g} := \{A, B\}_{\mathcal{L}}$ is the non-Abelian Lie algebra $\mathfrak{n}_2$. It is solvable: its derived subalgebra (Section B.2.1) is $\mathfrak{g}' = \mathbb{R}B$ and $\mathfrak{g}'' = 0$. Global coordinates of the second kind[30] must have the form

$$X(t) = e^{f(t)A}e^{g(t)B} \text{ with } f(0) = 0 = g(0).$$

$$\dot{X} = \dot{f}AX + \dot{g}e^{fA}Be^{-fA}X;\ \exp(f\,\mathrm{ad}_A)(B) = B,\ \text{so } uA + vB = \dot{f}A + \dot{g}e^{f}B;$$

$$f(t) = \int_0^t u(s)\,ds,\ g(t) = \int_0^t v(s)e^{-f(s)}\,ds. \qquad \triangle$$

---

[28] Wei and Norman [285], on the usefulness of Lie algebras in quantum physics, was very influential in bilinear control theory; for related operator calculus, see [195, 273, 286] and the CBH Theorem 2.1.

[29] Formulas for the ϕ_i in (2.37) were given also by Huillet et al. [134]. (The title refers to a minimal product formula, not input–output realization).

[30] The ansatz for $U(t)$ in [284] was misprinted.

Example 2.24. When the Lie algebra $\{B_1, \dots, B_m\}_{\mathcal{L}}$ is nilpotent one can get a finite and easily evaluated product representation for transition matrices of (2.2). Let

$$\dot{X} = (u_1 B_1 + u_2 B_2)X \text{ with } [B_1, B_2] = B_3 \neq 0,\ [B_1, B_3] = 0 = [B_2, B_3].$$

$$\begin{gathered}
\text{Set } \ X(t; u) := \mathrm{e}^{\phi_1(t)B_1}\mathrm{e}^{\phi_2(t)B_2}\mathrm{e}^{\phi_3(t)B_3};\\
\dot{X} = \dot{\phi}_1 B_1 X + \dot{\phi}_2 (B_2 + tB_3)X + \dot{\phi}_3 B_3 X;\\
\dot{\phi}_1 = u_1,\ \dot{\phi}_2 = u_2,\ \dot{\phi}_3 = -tu_2;\\
\phi_1(t) = \int_0^t u_1(s)\,ds,\ \phi_2(t) = \int_0^t u_2(s)\,ds,\ \phi_3(t) = \int_0^t \int_0^s u_2(r)\,dr\,ds
\end{gathered}$$

(using integration by parts). △

Example 2.25 (Wei and Norman [285]). Let $\dot{X} = (u_1 H + u_2 E + u_3 F)X$ with $[E, F] = H$, $[E, F] = 2E$ and $[F, H] = -2F$, so $\ell = 3$. Using (2.37) and Proposition 2.2 one obtains differential equations that have solutions for small $t \geq 0$:

$$\begin{gathered}
\dot{X} = \dot{\phi}_1 HX + \dot{\phi}_2 \mathrm{e}^{\phi_1 \operatorname{ad}_H}(E)X + \dot{\phi}_3 \mathrm{e}^{\phi_1 \operatorname{ad}_H}(\mathrm{e}^{\phi_2 \operatorname{ad}_E}(F))X;\\
\mathrm{e}^{\phi_1 \operatorname{ad}_H}(E) = \mathrm{e}^{-2\phi_1}E,\ \mathrm{e}^{\phi_1 \operatorname{ad}_H}(\mathrm{e}^{\phi_2 \operatorname{ad}_E}(F)) = \mathrm{e}^{2\phi_1}F + \phi_2 H + \phi_2^2 \mathrm{e}^{-2\phi_1}E;\ \text{so}\\
\begin{bmatrix}\dot{\phi}_1\\ \dot{\phi}_2\\ \dot{\phi}_3\end{bmatrix} = \begin{bmatrix}1 & 0 & \phi_2\\ 0 & \mathrm{e}^{-2\phi_1} & \phi_2^2\mathrm{e}^{-2\phi_1}\\ 0 & 0 & \mathrm{e}^{2\phi_1}\end{bmatrix}^{-1}\begin{bmatrix}u_1\\ u_2\\ u_3\end{bmatrix} \equiv \begin{bmatrix}1 & 0 & -\phi_2\mathrm{e}^{-2\phi_1}\\ 0 & \mathrm{e}^{2\phi_1} & -\phi_2^2\mathrm{e}^{-2\phi_1}\\ 0 & 0 & \mathrm{e}^{-2\phi_1}\end{bmatrix}\begin{bmatrix}u_1\\ u_2\\ u_3\end{bmatrix}.
\end{gathered}$$

△

Example 2.26 (Rotation group, IV). Using only the linear independence of the B_i and the relations $[B_1, B_2] = B_3$, $[B_2, B_3] = B_1$, and $[B_3, B_1] = B_2$, a product-form transition matrix for

$$\begin{aligned}
&\dot{X} = u_1 B_1 X + u_2 B_2 X,\ X(0) = I \ \text{ is } \ X(t; u) := \mathrm{e}^{\phi_1 B_1}\mathrm{e}^{\phi_2 B_2}\mathrm{e}^{\phi_3 B_3};\\
&u_1 B_1 + u_2 B_2 = \dot{\phi}_1 B_1 + \dot{\phi}_2\Big(\cos(\phi_1)B_2 + \sin(\phi_1)B_3\Big) +\\
&\qquad\qquad \dot{\phi}_3\Big(\cos(\phi_1)\cos(\phi_2)B_3 + \cos(\phi_1)\sin(\phi_2)B_1 - \sin(\phi_1)B_2\Big).\\
&\begin{bmatrix}1 & 0 & \cos(\phi_1)\sin(\phi_2)\\ 0 & \cos(\phi_1) & -\sin(\phi_1)\\ 0 & \sin(\phi_1) & \cos(\phi_1)\cos(\phi_2)\end{bmatrix}\begin{bmatrix}\dot{\phi}_1\\ \dot{\phi}_2\\ \dot{\phi}_3\end{bmatrix} = \begin{bmatrix}u_1\\ u_2\\ 0\end{bmatrix}.
\end{aligned}$$

△

Exercise 2.6. Use the basis $\{B_i\}_1^6$ given in Example 2.21 to solve (globally) $\dot{X} = \sum_{i=1}^6 u_i B_i X, X(0) = I$ as a product $X(t; u) = \exp(\phi_1(t)) \cdots \exp(\phi_6(t))$ by obtaining expressions for $\dot{\phi}_i$ in terms of the u_i and the ϕ_i. △

2.10 Complex Bilinear Systems*

Bilinear systems with complex state were not studied until recent years; see Section 6.5 for their applications in quantum physics and chemistry. Their algebraic geometry is easier to deal with than in the real case, although one loses the ability to draw insightful pictures.

2.10.1 *Special Unitary Group*

In the quantum system literature, there are several current mathematical models. In the model to be discussed here and in Section 6.5, the state is an operator (matrix) evolving on the special unitary group

$$\mathrm{SU}(n) := \{X \in \mathbb{C}^{n\times n} | \, X^*X = I, \det(X) = 1\}; \text{ it has the Lie algebra}$$
$$\mathfrak{su}(n) = \{B \in \mathbb{C}^{n\times n} | \, B + B^* = 0, \operatorname{tr} B = 0\}.$$

There is an extensive literature on unitary representations of semisimple Lie groups that can be brought to bear on control problems; see D'Alessandro [71].

Example 2.27. A basis over $\mathbb{R}$ for the Lie algebra $\mathfrak{su}(2)$ is given by[31] the skew-Hermitian matrices

$$P_x := \sqrt{-1}\begin{bmatrix}0&1\\1&0\end{bmatrix}, P_y := \begin{bmatrix}0&1\\-1&0\end{bmatrix}, P_z := \sqrt{-1}\begin{bmatrix}1&0\\0&-1\end{bmatrix};$$
$$[P_y, P_x] = 2P_z, [P_x, P_z] = 2P_y, [P_z, P_y] = 2P_x;$$
$$\text{if } U \in \boldsymbol{SU}(2),\ U = \exp(\alpha P_z)\exp(\beta P_y)\exp(\gamma P_z),\ \alpha, \beta, \gamma \in \mathbb{R}.$$

Using the $\boldsymbol{\alpha}$ representation, SU(2) has a real form that acts on $\mathbb{R}^4_*$; its orbits are 3-spheres. A basis over $\mathbb{R}$ of $\boldsymbol{\alpha}(\mathfrak{su}(2)) \subset \mathfrak{gl}(4, \mathbb{R})$ is

$$\left\{\begin{bmatrix}0&0&1&0\\0&0&0&-1\\-1&0&0&0\\0&1&0&0\end{bmatrix}, \begin{bmatrix}0&0&0&1\\0&0&1&0\\0&-1&0&0\\-1&0&0&0\end{bmatrix}, \begin{bmatrix}0&1&0&0\\-1&0&0&0\\0&0&0&1\\0&0&-1&0\end{bmatrix}\right\}. \qquad \triangle$$

[31] Physics textbooks such as Schiff [235] often use as basis for $\mathfrak{su}(2)$ the Pauli matrices $\sigma_x := \begin{bmatrix}0&1\\1&0\end{bmatrix}$, $\sigma_y := \begin{bmatrix}0&-i\\i&0\end{bmatrix}$, $\sigma_z := \begin{bmatrix}1&0\\0&-1\end{bmatrix}$.

2.10.2 Prehomogeneous Vector Spaces

A notion closely related to weak transitivity arose when analysts studied actions of complex groups on complex linear spaces. A prehomogeneous vector space[32] over $\mathbb{C}$ is a triple $\{\mathbf{G}, \varrho, \mathbf{L}\}$ where $\mathbf{L}$ is a finite-dimensional vector space over $\mathbb{C}$ or some other algebraically closed field; typically $\mathbf{L} = \mathbb{C}^n$. The group $\mathbf{G}$ is a connected reductive algebraic group with a realization (see page 241 in Section B.5.1) $\varrho : \mathbf{G} \to \mathrm{GL}(n, \mathbb{C})$; and there is an orbit S of the group action such that the Zariski closure of S (see Section C.2.3) is $\mathbf{L}$; $\mathbf{L}$ is "almost" a homogeneous space for $\mathbf{G}$, so is called a *prehomogeneous space*. Further, $\mathbf{G}$ has at least one relatively invariant polynomial P on $\mathbf{L}$, and a corresponding invariant affine variety V. A simple example of a prehomogeneous space is $\{\mathbb{C}_*, \varrho, \mathbb{C}\}$ with $\varrho : \ (\zeta_1, z_1) \mapsto \zeta_1 z_1$ and invariant variety $\mathrm{V} = \{0\}$. These triples $\{\mathbf{G}, \varrho, \mathbf{L}\}$ had their origin in the observation that the Fourier transform of a complex power of $P(z)$ is proportional to a complex power of a polynomial $\widehat{P}(s)$ on the Fourier-dual space; that fact leads to functional equations for multivariable zeta-functions (analogous to Riemann's zeta)

$$\zeta(P, s) = \sum_{x \in \mathbb{Z}^n \setminus \{0\}} \frac{1}{|P(x)|^s}.$$

For details see Kimura's book [159].

Corresponding to ρ, let $\varrho_{\mathfrak{g}}$ be the realization of $\mathfrak{g}$, the Lie algebra of $\mathbf{G}$; if there is a point $\xi \in \mathbf{L}$ at which the rank condition $\varrho_{\mathfrak{g}}(\xi) = \dim \mathbf{L}$ is satisfied, then there exists an open orbit of $\mathbf{G}$ and $\{\mathbf{G}, \varrho, \mathbf{L}\}$ is a prehomogeneous vector space.

Typical examples use two-sided group actions on $\mathbb{F}^n$. Kimura's first example has the action $X \mapsto AXB^\top$ where $X \in \mathbb{C}^{n \times n}$, $\mathbf{G} = \mathbf{H} \times \mathrm{GL}(n, \mathbb{R})$, A is in an arbitrary semisimple algebraic group $\mathbf{H}$, and $B \in \mathrm{GL}(n, \mathbb{R})$. The singular set is given by the vanishing of the nth-degree polynomial $\det(X)$, which is a relative invariant. Kimura states that the isotropy subgroup at $I_n \in C^{n \times n}$ is isomorphic to $\mathbf{H}$; that is because for $(Q, Q^\top)$ in the diagonal of $\mathbf{H} \times \mathbf{H}^\top$, one has $QI_nQ^{-1} = I_n$.

Such examples can be represented as linear actions by our usual method, $X \mapsto X^\flat$ with the action of $\mathbf{G}$ on $\mathbb{C}^\nu$ represented using Kronecker products. For that and other reasons, prehomogeneous vector spaces are of interest in the study of bilinear systems on complex groups and homogeneous spaces. Such a transfer of mathematical technology has not begun. The real matrix groups in this chapter, transitive or weakly transitive, need not be algebraic and their relation to the theory of orbits of algebraic groups is obscure. Rubenthaler [226] has defined and classified real prehomogeneous spaces

[32] Prehomogeneous vector spaces were introduced by Mikio Sato; see Sato and Shintani [234] and Kimura [159], on which I draw.

$\{\mathbf{G}, \varrho, \mathbb{R}^n\}$ in the case that V is a hypersurface and $\mathbf{G}$ is of parabolic type, which means that it contains a maximal solvable subgroup.

Example 2.28. Let ρ be the usual linear representation of a Lie group by matrices acting on $\mathbb{R}^3$. There are many Lie algebras for which $(\mathbf{G}, \rho, R^3)$ is a real prehomogeneous vector space; they have generic rank $\mathrm{r}(\mathfrak{g}) = 3$. Three Lie algebras are transitive on $\mathbb{R}^3_*$:

$$\mathfrak{gl}(3,\mathbb{R}): \ \ell = 9, \mathcal{P}(x) = 1; \quad \mathfrak{sl}(3,\mathbb{R}): \ \ell = 8, \mathcal{P}(x) = 1; \text{ and}$$
$$\mathfrak{so}(3) \oplus \mathbb{R}: \ \ell = 4, \mathcal{P}(x) = x_1^2 + x_2^2 + x_3^2.$$

There are many weakly transitive but nontransitive Lie subalgebras of $\mathfrak{gl}(3,\mathbb{R})$, for instance $\mathfrak{so}(2,1) \oplus \mathbb{R}$, whose singular locus V_0 is a quadric cone, and $\mathfrak{d}(3,\mathbb{R})$.

△

2.11 Generic Generation*

It is well known that the set of pairs of real matrices whose Lie algebra $\{A, B\}_{\mathcal{L}}$ is transitive on $\mathbb{R}^n_*$ is open and dense in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ [32]. The reason is that the set of entries $a_{i.j}, b_{i,j}$ for which $\ell < n^2$ is determined by the vanishing of polynomials in those entries; that is clear from `LieTree`. Here is a more subtle question:

Given numerically uncertain matrices A and B *known to belong to one of the transitive Lie algebras* $\mathfrak{g}$, are there nearby matrices $\tilde{A}, \tilde{B}$ in $\mathfrak{g}$ that generate all of $\mathfrak{g}$?

Corollary 2.3 will answer that question for all transitive Lie subalgebras of $\mathfrak{gl}(n,\mathbb{R})$.

Theorem 2.3 (Generic Transitivity). *In an arbitrarily small neighborhood of any given pair A, B of $n \times n$ matrices, one can find another pair of matrices that generate $\mathfrak{gl}(n,\mathbb{R})$.*

Proof.[33] Construct two matrices P, Q such that $\{P, Q\}_{\mathcal{L}} = \mathfrak{gl}(n,\mathbb{R})$. Let $P = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ be a real diagonal matrix satisfying

$$\alpha_i - \alpha_j \neq \alpha_k - \alpha_l \ \text{ if } (i, j) \neq (k, l), \tag{2.39}$$
$$\mathrm{tr}\, P \neq 0. \tag{2.40}$$

The set of real $n \times n$ matrices with zero elements on the main diagonal will be called $\mathcal{D}_0(n,\mathbb{R})$. Let Q be any matrix in $\mathcal{D}_0(n,\mathbb{R})$ such that all its off-diagonal

[33] A sketch of a proof is given in Boothby and Wilson [32, pp. 2,13]. The the proof given here is adapted from one in Agrachev and Liberzon [1]. Among other virtues, it provides most of the proof of Corollary 2.3.

elements are nonzero. In particular, we can take all the off-diagonal elements equal to 1. Suitable examples for $n = 3$ are

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}, \; Q = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

It is easy to check that if $i \neq j$, then $[P, 1^{i,j}] = (\alpha_i - \alpha_j)1^{i,j}$; so, using (2.39), Q does not belong to any ad_P-invariant proper subspace of $\mathcal{D}_0(n, \mathbb{R})$. If we construct the set

$$\mathcal{Q} := \{Q, \mathrm{ad}_P(Q), \ldots, \mathrm{ad}_P^{n(n-1)-1}(Q)\} \text{ we see } \operatorname{span} \mathcal{Q} = \mathcal{D}_0(n, \mathbb{R});$$

that is, each of the elementary matrices $1^{i,j}$, $i \neq j$ can be written as a linear combination of the $n(n-1)$ matrices in $\mathcal{Q}$. From brackets of the form $[1^{i,j}, 1^{j,i}]$, $i < j$, we can generate a basis $\{E_1, \ldots, E_{n-1}\}$ for the subspace of the diagonal matrices whose trace is zero. Since P has a nonzero trace by (2.40), we conclude that

$$\{P, Q, \mathrm{ad}_P(Q), \ldots, \mathrm{ad}_P^{n(n-1)-1}(Q), \; E_1, \ldots, E_{n-1}\}$$

is a basis for $\mathfrak{gl}(n, \mathbb{R})$, so $\{P, Q\}_{\mathcal{L}} = \mathfrak{gl}(n, \mathbb{R})$.

Second, construct the desired perturbations of A and B, the given pair of matrices. Using the matrices P and Q just constructed, define $A_s := A + sP$ and $B_s := B + sQ$, where $s \geq 0$.

Now construct a linearly independent list of matrices $\mathfrak{B}$ consisting of P, Q and $\nu - 2$, other Lie monomials generated from P, Q (as in Section 2.2.4, using the `LieWord` encoding to associate formal monomials with matrices). Then $\mathfrak{B}$ is a basis for $\mathfrak{gl}(n, \mathbb{R})$. Replacing P, Q in each of these ν Lie monomials by A_s, B_s respectively results in a new list $\mathfrak{B}_s$; and writing the coordinates of the ν matrices in $\mathfrak{B}_s$ relative to basis $\mathfrak{B}$, we obtain a linear coordinate transformation $T_s : \mathbb{R}^\nu \to \mathbb{R}^\nu$ parametrized by s. The $\nu \times \nu$ Jacobian matrix of this transformation is denoted by $J(s)$. Its determinant $\det(J(s))$ is a polynomial in s. As $s \to \infty$, T_s approaches a multiple of the identity; since $\det(J(s)) \to \infty$, it is not the zero polynomial. Thus $J(s)$ is nonsingular for all but finitely many values of s; therefore, if we take s sufficiently small, $\mathfrak{B}_s$ is a basis for $\mathfrak{gl}(n, \mathbb{R})$. □

Theorem 2.4 (Kuranishi [170, Th. 6]). *If $\mathfrak{g}$ is a real semisimple matrix Lie algebra, there exist two matrices $A, B \in \mathfrak{g}$ such that $\{A, B\}_{\mathcal{L}} = \mathfrak{g}$.*

A brief proof of Theorem 2.4 is provided by Boothby [29][34] where it is used, with Lemma 2.1, to prove [29, Theorem C], which is paraphrased here as Theorem 2.5.

[34] Also see Jacobson [143, Exer. 8, p. 150].

Lemma 2.1 (Boothby [29]). *Let $\mathfrak{g}$ be a reductive Lie algebra with center of dimension $\leq k$. If the semisimple part of $\mathfrak{g}$ is generated by k elements, then $\mathfrak{g}$ is generated by k elements.*

Theorem 2.5 (Th. C of [29]). *Each of the Lie algebras transitive on $\mathbb{R}^n_*$ can be generated from two matrices.*

All of these transitive Lie algebras are listed in Appendix D; the best known is $\mathfrak{sl}(n,\mathbb{R})$. The proof of Theorem 2.5 from the lemma uses the fact (see (D.4)) that the dimension k of the center of a transitive Lie algebra is no more than 2.

Corollary 2.3. *Let $\mathfrak{g}$ be a Lie algebra transitive on $\mathbb{R}^n_*$. Any two matrices A, B in $\mathfrak{g}$ have arbitrarily near neighbors A_s and B_s in $\mathfrak{g}$ that generate $\mathfrak{g}$.*

Proof. The proof is the same as the last paragraph of the proof of Theorem 2.3, except that P, Q, A_s, and B_s all lie in $\mathfrak{g}$ instead of $\mathfrak{gl}(n,\mathbb{R})$. □

If the entries of matrices A, B are chosen from any continuous probability distribution, then $\{A, B\}_{\mathcal{L}}$ is transitive with probability one. For this reason, in looking for random bilinear systems with interesting invariant sets, use discrete probability distributions; for instance, take $a_{i,j}, b_{i,j} \in \{-1, 0, 1\}$ with probability 1/3 for each of $\{-1, 0, 1\}$.

2.12 Exercises

Exercise 2.7. There are useful generalizations of homogeneity and of Example 2.8. Find the relative invariants of the linear dynamical system $\dot{x}_i = \beta_i x_i,\ i \in 1, \dots, n$. △

Exercise 2.8. Change a single operation in the script of `LieTree` so that it produces a basis for the associative matrix algebra $\{A, B\}_{\mathfrak{A}} \subset \mathbb{R}^{n\times n}$, and rename it `AlgTree`. See Sections A.2–A.2.2. △

Exercise 2.9. Show that if A and B can be brought into *real* triangular form by the same similarity transformation (see Section B.2.1) then there exist real linear subspaces $\mathcal{V}_1 \subset \dots \subset \mathcal{V}_{n-1},\ \dim \mathcal{V}_k = k$ such that $X\mathcal{V}_k \subset \mathcal{V}_k$ for each $X \in \{A, B\}_{\mathcal{L}}$. △

Exercise 2.10. Exercise 2.9 does not show that a solvable Lie algebra cannot be transitive. For a counterexample on $\mathbb{R}^2_*$, let $A = I$ and $B = J$; $\{I, J\}_{\mathcal{L}} \simeq \mathbb{C}$ is solvable, but transitive. Find out what can be said about the transitivity of higher-dimensional bilinear systems with solvable Lie algebras. Compare the open problem in Liberzon [183]. △

Exercise 2.11 (Coset space). Verify that $\mathrm{SL}(2,\mathbb{R})/\mathfrak{T}(2,\mathbb{R}) \simeq \mathbb{R}^2_*$, where $\mathfrak{T}(2,\mathbb{R})$ is the upper triangular subgroup of $\mathrm{GL}^+(2,\mathbb{R})$. △

Exercise 2.12 (W. M. Goldman [106]).
Show that the three dimensional complex Abelian Lie group **G** acting on $\mathbb{C}^3$ corresponding to the Lie algebra

$$\mathfrak{g} = \left\{ \begin{bmatrix} u_1 + u_2 & 0 & u_3 \\ 0 & u_1 & u_2 \\ 0 & 0 & u_1 \end{bmatrix}, u \in \mathbb{C}^3 \right\}, \text{ namely } \mathbf{G} = \left\{ \begin{bmatrix} e^{p+q} & 0 & e^p r \\ 0 & e^p & e^p q \\ 0 & 0 & e^p \end{bmatrix} \right\}$$

is transitive on $\{z \in \mathbb{C}^3 | z_3 \neq 0\}$; **G** is closed and connected, but not an algebraic group: look at the semisimple subgroup[35] **H** defined by $p = r = 0$. The Zariski closure of **G** contains the semisimple part of **H**. △

Exercise 2.13. Suppose that $1, a, b$ are not rationally related (take $a = \sqrt{2}$, $b = \sqrt{3}$, for instance). Verify by symbolic computation that the matrix Lie algebra

$$\mathfrak{g} := \text{span} \begin{bmatrix} u_1 + u_3 & u_2 - u_4 & 0 & 0 & u_5 & -u_6 \\ -u_2 + u_4 & u_1 + u_3 & 0 & 0 & u_6 & u_5 \\ 0 & 0 & u_1 & au_2 & u_3 + (a-b)u_8 & -u_4 + (a-b)u_7 \\ 0 & 0 & -au_2 & u_1 & u_4 + (b-a)u_7 & u_3 + (a-b)u_8 \\ 0 & 0 & 0 & 0 & u_1 & bu_2 \\ 0 & 0 & 0 & 0 & -bu_2 & u_1 \end{bmatrix}$$

has $\mathcal{P}(x) = x_5^2 + x_6^2$ and therefore is weakly transitive on $\mathbb{R}^6_*$. Show that the corresponding matrix group **G** is not closed. What group is the closure[36] of **G** in $GL^+(6, \mathbb{R})$? △

Exercise 2.14 (Vehicle attitude control). If A, B are the matrices of Example 2.19, show that $\dot{x} = u(t)Ax + v(t)Bx$ on $\mathbb{R}^3$ is controllable on the unit sphere S^2. How many constant-control trajectories are needed to join $x(0) = (1, 0, 0)$ with any other point on S^2? △

Problem 2.2. How does the computational complexity of `LieTree` depend on m and n? △

Problem 2.3. For most of the Lie algebras $\mathfrak{g}$ transitive on $\mathbb{R}^n_*$ (Appendix D), the isotropy subalgebras $\mathfrak{h}_x \subset \mathfrak{g}$ are not known; for low dimensions, they can readily be calculated at $p = (1, 0, \ldots, 0)$. △

Problem 2.4. Implement the algorithms in Boothby and Wilson [32] that decide whether a generated matrix Lie algebra $\mathbf{B}^m_{\mathcal{L}}$ is in the list of transitive Lie algebras in Section D.2. △

---

[35] For any element g in an algebraic subgroup **H**, both the semisimple part g_s and unipotent part g_u must lie in **H**. See Humphreys [135, Theorem 1.15].

[36] Hint: The group $\mathbf{G}^c$ that is the closure of **G** must contain $\exp(u_9 B)$ where B is the 6×6 matrix whose elements are zero except for $B_{34} = 1$, $B_{43} = -1$. Is $\mathbf{G}^c$ transitive on $\mathbb{R}^6_*$?

Chapter 3
Systems with Drift

3.1 Introduction

This chapter is concerned with the stabilization[1] and control of bilinear control systems with drift term $A \neq 0$ and constraint set $\Omega \subset \mathbb{R}^m$:

$$\dot{x} = Ax + \sum_{i=1}^{m} u_i(t)B_i x, \ x \in \mathbb{R}^n, \ u(\cdot) \in \Omega, \ t \geq 0; \tag{3.1}$$

as in Section 1.5, this system's transition matrices $X(t; u)$ satisfy

$$\dot{X} = AX + \sum_{1}^{m} u_i(t)B_i X, \ X(0) = I_n, \ u(\cdot) \in \Omega, \ t \geq 0. \tag{3.2}$$

The control system (3.1) is stabilizable if there exists a feedback control law $\phi : \mathbb{R}^n \to \Omega$ and an open basin of stability $\mathbf{U}$ such that the origin is an asymptotically stable equilibrium for the dynamical system

$$\dot{x} = Ax + \sum_{i=1}^{m} \phi_i(x)B_i x, \ \xi \in \mathbf{U}. \tag{3.3}$$

It is not always possible to fully describe the basin of stability, but if $\mathbf{U} = \mathbb{R}^n_*$, (3.1) is said to be globally stabilizable; if $\mathbf{U}$ is bounded, system (3.1) is called locally stabilizable.[2]

---

[1] For the stabilization of a nonlinear system $\dot{x} = f(x) + ug(x)$ with output $y = h(x)$ by output feedback $u = \phi \circ h(x)$, see Isidori [140, Ch. 7]. Most works on such problems have the essential hypothesis $g(0) \neq 0$, which cannot be applied to (3.1) but is applicable to the biaffine systems in Section 3.8.3.

[2] Sometimes there is a stronger stability requirement to be met, that $\|x(t)\| \to 0$ faster than $e^{-t\epsilon}$; one can satisfy this goal to replace A in (3.3) with $A + \epsilon I$ and find ϕ that will stabilize the resulting system.

DOI 10.1007/978-1-4020-9613-6_3,

As in Chapter 1, $\mathbf{S}_\Omega$ denotes the semigroup of matrices $X(t;u)$ that satisfy (3.2); the set $\mathbf{S}_\Omega\xi$ is called the orbit of the semigroup through ξ. For system (3.1) controllability (on $\mathbb{R}^n_*$, see Definition 1.9 on page 21) is equivalent to the following statement:

$$\text{For all } \xi \in \mathbb{R}^n_*,\ \mathbf{S}_\Omega\xi = \mathbb{R}^n_*. \tag{3.4}$$

The controllability problem is to find properties of (3.1) that are necessary or sufficient for (3.4) to hold. The case $n = 2$ of (3.1) provides the best results for this problem because algebraic computations are simpler in low dimension and we can take advantage of the topology of the plane, where curves are also hypersurfaces.

Although both stabilization and control of (3.1) are easier when $m > 1$, much of this chapter simplifies the statements and results by specializing to

$$\dot{x} = Ax + uBx,\ x(0) = \xi \in \mathbb{R}^n,\ \ u(t) \in \Omega. \tag{3.5}$$

Definition 3.1. For $A, B \in \mathbb{R}^{n\times n}$, define the relation $\approx$ on matrix pairs by

$$(A, B) \approx (P^{-1}(A + \mu B)P, P^{-1}BP) \text{ for some } \mu \in \mathbb{R} \text{ and } P \in \mathrm{GL}(n, \mathbb{R}). \qquad \triangle$$

Lemma 3.1.
(I) $\approx$ is an equivalence relation;
(II) if $\Omega = \mathbb{R}$, $\approx$ preserves both stabilizability and controllability for (3.5).

Proof. (I): The relation $\approx$ is the composition of two equivalence relations: the shift $(A, B) \sim_1 (A + \mu B, B)$ and the similarity $(A, B) \sim_2 (P^{-1}AP, P^{-1}BP)$.
(II): Asymptotic stability of the origin for $\dot{x} = Ax + \phi(x)Bx$ is not changed by linear changes of coordinates $x = Py$; $\dot{x} = (A + \mu B)x + uBx$ is stabilizable by $u = \phi(x) - \mu$. Controllability is evidently unaffected by $u \mapsto u + \mu$. Let $\tilde{\mathbf{S}}$ be the transition semigroup for $\dot{y} = P^{-1}APy + uP^{-1}BPy$. Using the identity $\exp(P^{-1}XP) = P^{-1}\exp(X)P$ and remembering that the elements of $\mathbf{S}$ are products of factors $\exp(t(A+uB))$, the statement $\mathbf{S}\xi = \mathbb{R}^n_*$ implies $\tilde{\mathbf{S}}P\xi = P\mathbb{R}^n{*} = \mathbb{R}^n_*$, so $\sim_2$ preserves controllability. □

Contents of this Chapter

As a gentle introduction to stabilizability for bilinear systems, Section 3.2 introduces stabilization by constant feedback. Only Chapter 1 is needed to read that section. It contains necessary and sufficient conditions of Chabour et al. [52] for stabilization of planar systems with constant controls.

Section 3.3 gives some simple facts about controllability and noncontrollability, the geometry of attainable sets, and some easy theorems. To obtain accessibility conditions for (3.1), one needs the ad-condition, which is stronger than the Lie algebra rank condition; it is developed in Section 3.4, to which Chapter 2 is prerequisite. With this condition, Section 3.5 can deal with controllability criteria given indefinitely small bounds on control amplitude; other applications of the ad-condition lead-in to the stabilizing feedback laws of Jurdjevic and Quinn [149] and Chabour et al. [52]. Sometimes linear feedback laws lead to quadratic dynamical systems that although not asymptotically stable, nevertheless have global attractors that live in small neighborhoods of the origin and may be chaotic; see Section 3.6.5.

One approach to controllability of (3.1) is to study the semigroup $\mathbf{S}_\Omega$ generated by (3.2). If $\mathbf{S}_\Omega$ is actually a Lie group $\mathbf{G}$ known to be transitive on $\mathbb{R}^n_*$, then (3.1) is controllable. This method originated with the work of Jurdjevic and Kupka [148], where $\mathbf{G} = \mathrm{SL}(n, \mathbb{R})$, and is surveyed in Section 3.7. Alongside that approach (and beginning to overlap) is a recent research on subsemigroups of Lie groups; see Section 3.7.2.

Controllability and stabilization for biaffine systems are briefly discussed in Section 3.8, as well as an algebraic criterion for their representation as linear control systems. A few exercises are given in Section 3.9.

3.2 Stabilization with Constant Control

Here we consider a special stabilization problem: Given matrices A, B, for what real constants $\mu \in \Omega$ is $A + \mu B$ a Hurwitz matrix?

Its characteristic polynomial is

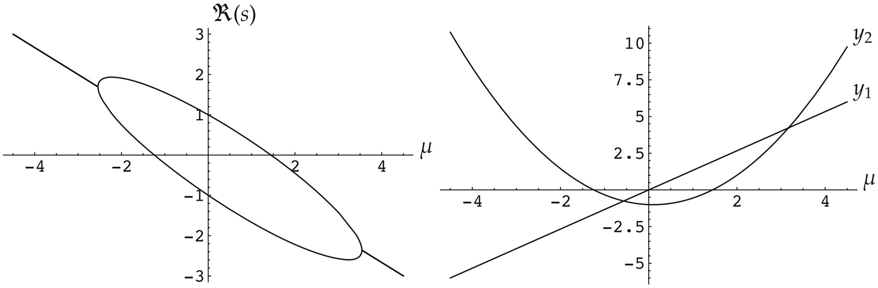
$$\boldsymbol{p}_{A+\mu B}(s) := \det(sI - A - \mu B) = \prod_{i=1}^{n}(s - \lambda_i(\mu)),$$

where the λ_i are algebraic functions of μ. Let $\rho_i(\mu) := \Re(\lambda_i(\mu))$. One can plot the ρ_i to find the set (possibly empty) $\{\mu | \max_i \rho_i(\mu) < 0\}$ which answers our question. The Mathematica script in Fig. 3.1 calculates the $\rho_i(\mu)$ and generates some entertaining plots of this kind; an example with $n = 2$ is part (a) of the figure. The polynomial $\boldsymbol{p}_{A+\mu B}(s)$ is a Hurwitz[3] polynomial (in s) when all the ρ_i values lie below the μ axis.

3.2.1 *Planar Systems*

In the special case $n = 2$, the characteristic polynomial of $A + \mu B$ is

[3] A polynomial $p(s)$ over $\mathbb{F}$ is called a Hurwitz polynomial if all of its roots have negative real parts.



(a) (b)

```
n=2; r=4.5; A={{-1,0},{1/3,1}}; B={{-2/3,1/3},{-1/3,-2/3}};
For[i=1,i<=n,i++,eigr[i][u_]:=Re[Eigenvalues[A+u*B][[i]]]];      Y[u_] :=
Table[{u, eigr[i][u]}, {i, n}]; ParametricPlot[Evaluate[Y[u]],{u, -r, r}];
P[s_,u_]:= Det[s* Id-(A+u* B)]; y1[u_]:=Coefficient[P[s,u],s]; y2[u_]:=P[0,u];
Plot[{y1[u], y2[u]},{u,-r,r}]}]
```

Fig. 3.1. Mathematica script with plots (**a**) $\Re\,(\mathrm{spec}(A+\mu B))$ vs. μ and (**b**) the coefficients $y_1(\mu), y_2(\mu)$ of $\boldsymbol{p}(s)$, which is a Hurwitz polynomial for $\mu > 1.45$.

$$\boldsymbol{p}_{A+\mu B}(s) = s^2 + s y_1(\mu) + y_2(\mu) \quad \text{where} \quad y_1(\mu) := -\operatorname{tr}(A) - \mu \operatorname{tr}(B),$$
$$y_2(\mu) := \det(A) + \mu\Big(\operatorname{tr}(A)\operatorname{tr}(B) - \operatorname{tr}(AB)\Big) + \mu^2 \det(B).$$

The coefficients y_1, y_2 are easily evaluated and plotted as in Fig. 3.1(b) which was produced by the Mathematica commands shown in the figure.

Proposition 3.1. *On $\mathbb{R}^2$, the necessary and sufficient condition for the dynamical system $\dot{x} = Ax + \mu Bx$ to be globally asymptotically stable for a constant $\mu \in \Omega$ is $\{\mu |\, y_1(\mu) > 0\} \cap \{\mu |\, y_2(\mu) > 0\} \cap \Omega \neq \emptyset$.*

Proof. The positivity of both y_1 and y_2 is necessary and sufficient for $s^2 + s y_1(\mu) + y_2(\mu)$ to be a Hurwitz polynomial. □

Finding the two positivity sets on the μ axis given the coefficients a_i, b_i is easy with the guidance of a graph like that of Fig. 3.1(b) to establish the relative positions of the two curves and the abscissa; this method is recommended if Ω is bounded and feasible values of μ are needed. Plots of the real parts of eigenvalues can be used for systems of higher dimension and $m > 1$. I do not know of any publications about stabilization with constant feedback for multiple inputs.

If the controls are unrestricted, we see that if $\operatorname{tr}(B) \neq 0$ and $\det(B) > 0$ there exists a constant feedback μ for which $y_1(\mu)$ and $y_2(\mu)$ are both positive; then $A + \mu B$ is a Hurwitz matrix. Other criteria are more subtle. For instance, Proposition 3.1 can (if $\Omega = \mathbb{R}$) be replaced by the necessary and sufficient conditions given in Bacciotti and Boieri [14] for constant feedback.

A systematic and self-contained discussion of stabilization by feedback is found in Chabour et al. [52]; our brief account of this work uses the equivalence relations $\sim_1$, $\sim_2$, and $\approx$ of Lemma 3.1. The intent in [52] is to find those systems (3.5) on $\mathbb{R}^2_*$ that permit stabilizing laws. These systems must, from Lemma 3.1, fall into classes equivalent under $\approx$.

The invariants of a matrix pair (A, B) under similarity ($\sim_2$) are polynomials in $a_1, \dots, b_2$, the coefficients of the parametrized characteristic polynomial

$$\boldsymbol{p}_{A,B}(s,t) := \det(sI - A - tB) = s^2 + (a_1 + a_2 t)s + (b_0 + b_1 t + b_2 t^2);$$

they are discussed in Section A.4. Some of them are (Exercise 3.1) also invariant under $\sim_1$ and will be used in this section and in Section 3.6.3 to sketch the work of [52]. These $\approx$-invariants of low degree are

$$\begin{aligned}
&a_2 = -\operatorname{tr}(B), \quad b_2 = \det(B) = \frac{1}{2}(\operatorname{tr}(B)^2 - \operatorname{tr}(B^2)),\\
&\mathbf{p}_1 := \Big(\operatorname{tr}(A)\operatorname{tr}(B) - \operatorname{tr}(AB)\Big)^2 - 4\det A \det B = b_1^2 - 4b_0 b_2,\\
&\mathbf{p}_2 := \operatorname{tr}(B)\operatorname{tr}(AB) - \operatorname{tr}(A)\operatorname{tr}(B^2) = 2a_1 b_2 - a_2 b_1,\\
&\mathbf{p}_3 := 2\operatorname{tr}(A)\operatorname{tr}(B)\operatorname{tr}(AB) - \operatorname{tr}(A^2)\operatorname{tr}(B)^2 - \operatorname{tr}(A)^2\operatorname{tr}(B^2)\\
&\quad\;\; = 2b_0 a_2^2 - 2a_1 a_2 b_1 + 2a_1^2 b_2,\\
&\mathbf{p}_4 := \chi(A,A)\chi(B,B) - \chi(A,B)^2\\
&\quad\;\; = b_0 a_2^2 - a_1 a_2 b_1 + b_1^2 - 4b_0 b_2 + a_1^2 b_2 = \Delta/16
\end{aligned}$$

where[4] $\chi(X,Y) := 2\operatorname{tr}(XY) - \operatorname{tr}(X)\operatorname{tr}(Y)$ and Δ is the invariant defined by (A.18) on page 224.

The following is a summary of the constant-control portions of Theorems 2.1.1, 2.2.1, and 2.3.1 in [52]. For each real canonical form of B, the proof in [52] shows that the stated condition is equivalent to the existence of some μ such that the characteristic polynomial's coefficients $y_1(\mu) = a_1 + a_2\mu$ and $y_2(\mu) = b_0 + b_1\mu + b_2\mu^2$ are both positive.

Theorem 3.1 ([52]). *The system $\dot{x} = Ax + \mu Bx$ on $\mathbb{R}^2_*$ is globally asymptotically stable for some constant μ if and only if one of the following conditions is met:*
***a**. B is similar to a real diagonal matrix and one of these is true:*

> ***a**.1* $\det(B) > 0$,
> ***a**.2* $\mathbf{p}_1 > 0$ *and* $\mathbf{p}_2 > 0$,
> ***a**.3* $\mathbf{p}_1 > 0$ *and* $\mathbf{p}_3 > 0$,
> ***a**.4* $\det(B) = 0, \mathbf{p}_1 = 0$ *and* $\mathbf{p}_3 > 0$;

***b**. B is similar to* $\left(\begin{smallmatrix}\beta & 1\\ 0 & \beta\end{smallmatrix}\right)$ *and one of these is true:*

---

[4] As a curiosity, [52] remarks that for $\mathfrak{gl}(2,\mathbb{R})$ the Cartan–Killing form is $\operatorname{tr}(\operatorname{ad}_X \operatorname{ad}_Y) = 2\chi(X,Y)$; that can be seen from (A.12).

$\boldsymbol{b}.1\ \operatorname{tr}(B) \neq 0,$
$\boldsymbol{b}.2\ \operatorname{tr}(B) = 0$ *and* $\operatorname{tr}(AB) \neq 0$ *and* $\operatorname{tr}(A) < 0,$
$\boldsymbol{b}.3\ \operatorname{tr}(B) = 0$ *and* $\operatorname{tr}(AB) = 0$ *and* $\operatorname{tr}(A) < 0$ *and* $\det(A) > 0.$

***c**. B has no real eigenvalues and one of these is true:*

$\boldsymbol{c}.1\ \operatorname{tr}(B) \neq 0,$
$\boldsymbol{c}.2\ \operatorname{tr}(B) = 0$ *and* $\operatorname{tr}(A) < 0.$

Exercise 3.1. The invariance under $\sim_1$ of $\operatorname{tr}(B)$ and $\det(B)$ is obvious; they are the same when evaluated for the corresponding coefficients of $\det(sI - A - (t + \mu)B)$. Verify the invariance of $\mathbf{p}_1, \ldots, \mathbf{p}_4$. △

3.2.2 *Larger Dimensions*

For $n > 2$, there are a few interesting special cases; one was examined in Luesink and Nijmeijer [194]. Suppose that $\{A, B\}_{\mathcal{L}}$ is solvable (Section B.2.1). By a theorem of Lie [282, Th. 3.7.3], any solvable Lie algebra $\mathfrak{g}$ has an upper triangular representation in $\mathbb{C}^n$. It follows that there exists a matrix $S \in \mathbb{C}^n$ for which $\widetilde{A} = S^{-1}AS$ and $\widetilde{B} = S^{-1}BS$ are simultaneously upper triangular.[5] The diagonal elements of $\widetilde{A}$ and $\widetilde{B}$ are their eigenvalues, α_i and β_i, respectively. Similar matrices have the same spectra, so

$$\operatorname{spec}(\widetilde{A} + \mu\widetilde{B}) = \{\alpha_i + \mu\beta_i\}_{i=1}^{n} = \operatorname{spec}(A + \mu B).$$

If for some μ the inequalities

$$\Re(\alpha_i + \mu\beta_i) \leq -\epsilon < 0,\ i \in 1, \ldots, n$$

are satisfied, then $A + \mu B$ is a Hurwitz matrix and solutions of $\dot{x} = (A + \mu B)x$ decay faster than $e^{-\epsilon t}$. There are explicit algorithms in [194] that test for solvability and find the triangularizations, and an extension of this method to the case that the A, B are simultaneously conjugate to matrices in a special block-triangular form.

Proposition 3.2 (Kalouptsidis and Tsinias [155, Th. 4.1]).
If B or $-B$ is a Hurwitz matrix there exists a constant control u such that the origin is globally asymptotically stable for (3.5).

[5] There exists a hyperplane L in $\mathbb{C}^n$ invariant for solvable $\{A, B\}_{\mathcal{L}}$; if the eigenvalues of A, B are all real, their triangularizations are real and L is real. Also note that $\mathfrak{g}$ is solvable if and only if its Cartan–Killing form $\chi(X, Y)$ vanishes identically (Theorem B.4).

Proof. The idea of the proof is to look at $B + \frac{1}{u}A$ as a perturbation of B. If B is a Hurwitz matrix, there exists $P \gg 0$ such that $B^\top P + PB = -I$. Use the gauge function $V(x) = x^\top Px$ and let $Q := A^\top P + PA$.[6] Then $\dot{V}(x) = x^\top Qx - ux^\top x$. There exists $\kappa > 0$ such that $x^\top x \geq \kappa x^\top Px$ on $\mathbb{R}^n$. For any $\lambda < 0$, choose any $u > \|Q\| - 2(\lambda/k)$. Then by standard inequalities $\dot{V} \leq 2\lambda V$. If $-B$ is a Hurwitz matrix make the obvious changes. □

Example 3.1. Bilinear systems of second and third order, with an application to thermal fluid convection models, were studied in Čelikovský [49]. The systems studied there are called *rotated semistable* bilinear systems, defined by the requirements that in $\dot{x} = Ax + uBx$

$\boldsymbol{C_1}$ **:** $B^\top + B = 0$ (i.e., $B \in \mathfrak{so}(3)$) and
$\boldsymbol{C_2}$ **:** A has eigenvalues in both the negative and positive half-planes.
It is necessary, for a semistable systems to be stabilizable, that
$\boldsymbol{C_3}$ **:** $\det(A) \neq 0$ and $\operatorname{tr}(A) < 0$.

Čelikovský [49, Th. 2] gives conditions sufficient for the existence of constant stabilizers. Suppose that on $\mathbb{R}^3_*$ the system (3.5) satisfies the conditions $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$, and $\boldsymbol{C_3}$, and that $A = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$. Take

$$B = \begin{bmatrix} 0 & c_3 & -c_2 \\ -c_3 & 0 & c_1 \\ c_1 & -c_1 & 0 \end{bmatrix}, \gamma := \frac{\lambda_1 c_1^2 + \lambda_2 c_2^2 + \lambda_3 c_3^2}{c_1^2 + c_2^2 + c_3^2}.$$

The real parts of the roots are negative, giving a stabilizer, for any μ such that the three Routh–Hurwitz conditions for $n = 3$ are satisfied by $\det(sI - (A + \mu B))$: they are equivalent to

$$\operatorname{tr}(A) < 0, \quad \det(A) + \gamma\mu^2 < 0,$$
$$(\lambda_1 + \lambda_2)(\lambda_2 + \lambda_3)(\lambda_1 + \lambda_3) < (\gamma - \operatorname{tr}(A))\mu^2.$$

Theorem 3.9 below deals with planar rotated semistable systems. △

3.3 Controllability

This section is about controllability, the transitivity on $\mathbb{R}^n_*$ of the matrix semigroup $\mathbf{S}_\Omega$ generated by (3.2). In $\mathbb{R}^{n\times n}$, all of the orbits of (3.2) belong to $\mathrm{GL}^+(n, \mathbb{R})$ or to one of its proper Lie subgroups. $\mathbf{S}_\Omega$ is a subsemigroup of any closed matrix group $\mathbf{G}$ whose Lie algebra includes $A, B_1, \ldots, B_m$.

A long term area of research, still incomplete, has been to find conditions sufficient for the transitivity of $\mathbf{S}_\Omega$ on $\mathbb{R}^n_*$. The foundational work on this

[6] This corrects the typographical error "$V(x) = x^\top Qx$" in [155].

problem is Jurdjevic and Sussmann [151]; a recommended book is Jurdjevic [147]. Beginning with Jurdjevic and Kupka [148] the study of system (3.2)'s action on $\mathbf{G}$ has concentrated on the case that $\mathbf{G}$ is either of the transitive groups $\mathrm{GL}^+(n,\mathbb{R})$ or $\mathrm{SL}(n,\mathbb{R})$; the goal is to find sufficient conditions for $\mathbf{S}_\Omega$ to be $\mathbf{G}$. If so, then (3.1) is controllable on $\mathbb{R}^n_*$ (Section 3.7). Such sufficient conditions for more general semisimple $\mathbf{G}$ were given in Jurdjevic and Kupka [148].

3.3.1 Geometry of Attainable Sets

Renewing definitions and terminology from Chapter 1, as well as (3.2), the various attainable sets[7] for systems (3.1) on $\mathbb{R}^n_*$ are

$$\mathcal{A}_{t,\Omega}(\xi) := \{X(t;u)\xi \,|\, u(\cdot) \in \Omega\};$$

$$\mathcal{A}^T_\Omega(\xi) := \bigcup\nolimits_{0\le t\le T} \mathcal{A}_{t,\Omega}(\xi);$$
$$\mathcal{A}_\Omega(\xi) = \mathbf{S}_\Omega \xi.$$

If $\mathbf{S}_\Omega\xi = \mathbb{R}^n_*$ for all ξ, we say $\mathbf{S}_\Omega$ acts transitively on $\mathbb{R}^n_*$ and that (3.1) is controllable on $\mathbb{R}^n_*$. Different constraint sets yield different semigroups, the largest, $\mathbf{S}$, corresponding to $\Omega = \mathbb{R}^m$.

For clarity, some of the Lie algebras and Lie groups in this section are associated with their defining equations by using the tags of Equations (α) and (β), below, as subscripts.

A symmetric extension of system (3.1) can be obtained by assigning an input u_0 to A:

$$\dot{x} = (u_0 A + \sum_0^m u_i(t) B_i)x, \qquad (\alpha)$$

where (usually) $u_0 = \pm 1$ and A may or may not be linearly independent of the B_i. Using Chapter 2, there exists a Lie group of transition matrices of this system (closed or not) denoted by $\mathbf{G}_\alpha \subset \mathrm{GL}^+(n,\mathbb{R})$.

The `LieTree` algorithm can be used to compute a basis for the Lie algebra

$$\mathfrak{g}_\alpha := \{A, B_1, \dots, B_m\}_{\mathcal{L}}.$$

This Lie algebra corresponds to a matrix Lie subgroup $\mathbf{G}_\alpha \subset \mathrm{GL}^+(n,\mathbb{R})$; that $\mathbf{G}_\alpha$ is the group of transition matrices of (α) as shown in Chapter 2. Then $\mathbf{S}_\Omega$ is a subsemigroup of $\mathbf{G}_\alpha$.

[7] Attainable sets in $\mathbb{R}^{n\times n}$ can be defined in a similar way for the matrix system (3.2) but no special notation will be set up for them.

Proposition 3.3. *If the system* (3.1) *is controllable on* $\mathbb{R}^n_*$, *then* $\mathfrak{g}_\alpha$ *satisfies the LARC.*

Proof. Imposing the constraint $u_0 = 1$ in (α) cannot increase the content of any attainable set for (3.1), nor destroy invariance of any set; so a set $\mathbf{U} \in \mathbb{R}^n$ invariant for the action of $\mathbf{G}_\alpha$ is invariant for (3.1). Controllability of (3.1) implies that $\mathbf{G}_\alpha$ is transitive, so $\mathfrak{g}_\alpha$ is transitive. □

Proposition 3.4. *If* $\mathbf{S}_\Omega$ *is a matrix group then* $\mathbf{S}_\Omega = \mathbf{G}_\alpha$.

Proof. By its definition, $\mathbf{S}_\Omega$ is path-connected; by hypothesis, it is a matrix subgroup of $\mathrm{GL}(n, \mathbb{R})$, so from Theorem B.11 (Yamabe's Theorem) it is a Lie subgroup[8] and has a Lie algebra, which evidently is $\mathfrak{g}_\alpha$, so $\mathbf{S}_\Omega$ can be identified with $\mathbf{G}_\alpha$. □

Following Jurdjevic and Sussmann [151], we will define two important Lie subalgebras[9] of $\mathfrak{g}_\alpha$ that will be denoted by $\mathfrak{g}_\beta$ and $\mathfrak{i}_0$.

Omitting the drift term A from system (3.2), we obtain

$$\dot{X} = \sum_1^m u_i(t) B_i X, \ X(0) = I, \tag{β}$$

which generates a connected Lie subgroup $\mathbf{G}_\beta \subseteq \mathbf{G}_\alpha$ with corresponding Lie algebra $\mathfrak{g}_\beta := \{B_1, \ldots, B_m\}_{\mathcal{L}}$.

If $A \in \mathrm{span}\{B_1, \ldots, B_m\}$ and $\Omega = \mathbb{R}^m$ then the attainable sets in $\mathbb{R}^{n\times n}$ are identical for systems (3.2) and (β).

Proposition 3.5. *With* $\Omega = \mathbb{R}^m$, *the closure* $\overline{\mathbf{G}}_\beta$ *of* $\mathbf{G}_\beta$ *is a subset of* $\overline{\mathbf{S}}$, *the closure of* $\mathbf{S}$ *in* $\mathbf{G}_\alpha$; *it is either in the interior of* $\overline{\mathbf{S}}\xi$ *or it is part of the boundary.*

Proof. ([151, Lemma 6.4]). Applying Proposition 2.7, all the elements of $\overline{\mathbf{G}}_\beta$ are products of terms $\exp(tu_i B_i)$, $i \in 1, \ldots, m$. Fix t; for each $k \in \mathbb{N}$, define the input $u^{(k)}$ to have its ith component k and the other $m-1$ inputs zero. Then $X(t/k; u^{(k)}) = \exp(tA/k + tu_i B_i)$; take the limit as $k \to \infty$. (Example 1.7 illustrates the second conclusion.) □

The Lie algebra ideal in $\mathfrak{g}_\alpha$ generated by $\{B_i\}_1^m$ is[10]

$$\mathfrak{i}_0 := \{B_1, \ldots, B_m, [A, B_1], \ldots, [A, B_m]\}_{\mathcal{L}}; \tag{3.6}$$

[8] Jurdjevic and Sussmann [151, Th. 4.6] uses Yamabe's Theorem this way.

[9] Such Lie algebras were first defined for general real-analytic systems $\dot{x} = f(x) + ug(x)$ on manifolds M; [80, 251] and independently [269] showed that if the first homotopy group $\pi_1(M)$ contains no element of infinite order then controllability implies full Lie rank of $\mathfrak{i}_0$. The motivation of [80] came from Example 1.8 where $\pi_1(\mathbb{R}^2_*) = \mathbb{Z}$.

[10] This definition of the zero-time ideal $\mathfrak{i}_0$ differs only superficially from the definition in Jurdjevic [147, p. 59].

it is sometimes called the zero-time Lie algebra. Denote the codimension of $\mathfrak{i}_0$ in $\mathfrak{g}_\alpha$ by δ; it is either 0, if $A \in \mathfrak{i}_0$, or 1 otherwise. Let $\mathbf{H}$ be the connected Lie subgroup of $\mathbf{G}_\alpha$ corresponding to $\mathfrak{i}_0$. If $\delta = 0$ then $\mathbf{H} = \mathbf{G}_\alpha$. If $\delta = 1$, the quotient of $\mathfrak{g}_\alpha$ by $\mathfrak{i}_0$ is $\mathfrak{g}_\alpha/\mathfrak{i}_0 = \operatorname{span} A$. Because $\mathbf{H} = \alpha^{-1}\mathbf{H}\alpha$ for all $\alpha \in \mathbf{G}_\alpha$, $\mathbf{H}$ is a normal subgroup of $\mathbf{G}_\alpha$ and $\mathbf{G}_\alpha/\mathbf{H} = \exp(\mathbb{R}A)$. From that it follows that in $\mathbb{R}^n_*$ the set of states attainable at t for (3.2) satisfies

$$\mathcal{A}_{t,\Omega}(\xi) \subset \mathbf{H}e^{tA}\xi. \tag{3.7}$$

The group orbit $\mathbf{H}\xi$ is called the zero-time orbit for (3.1).

Let $\ell = \dim \mathfrak{g}_\alpha$ and find a basis $\{F_1, \dots, F_{\ell-\delta}\}$ for $\mathfrak{i}_0$ by applying the `LieTree` algorithm to the list $\{B_1, \dots, [A, B_m]\}$ in (3.6). The augmented rank matrix for $\mathfrak{i}_0$ is

$$M_0^+(x) := \begin{bmatrix} x \ F_1 x \ \cdots \ F_{\ell-\delta} x \end{bmatrix}; \tag{3.8}$$

its singular polynomial $\mathcal{P}_0^+(x)$ is the GCD of its $n \times n$ minors (or of their Gröbner basis). Note that one can test subsets of the minors to speed the computation of the GCD or the Gröbner basis.

3.3.2 *Traps*

It is useful to have examples of uncontrollability for (3.5). If the Lie algebra $\mathfrak{g} := \{A, B\}_\mathcal{L}$ is not transitive, from Proposition 3.3 invariant affine varieties for (3.1) of positive codimension can be found by the methods of Section 2.5.2, so assume the system has full Lie rank. In known examples of uncontrollability, the attainable sets are partially ordered in the sense that if $T > \tau$ then $\mathcal{A}^\tau_\Omega(\xi) \subsetneq \mathcal{A}^T_\Omega(\xi)$. Necessarily, in such examples, the ideal $\mathfrak{i}_0$ has codimension one in $\mathfrak{g}$.

Definition 3.2. A trap at ξ is an attainable set $\mathbf{S}_\Omega\xi \subsetneq \mathbb{R}^n_*$; so it is an invariant set for $\mathbf{S}_\Omega$. (A trap is called an invariant control set in Colonius and Kliemann [63] and other recent works.)

In our first examples, a Lyapunov-like method is used to find a trap. If there is a continuously differentiable gauge function $V \gg 0$ that grows along trajectories,

$$\dot{V}(x) := x^\top (A + uB)^\top \frac{\partial V}{\partial x} \geq 0 \text{ for all } u \in \Omega, x \in \mathbb{R}^n, \tag{3.9}$$

then the set $\{x \,|\, V(x) \geq V(\xi)\}$ is a trap at ξ.

It is convenient to use the vector fields

$$\mathbf{b} := x^\top B^\top \frac{\partial}{\partial x},$$

where $B \in \mathfrak{g}$. Let $\mathbf{i}_0 := \{\mathbf{b} | B \in \mathfrak{i}_0\}$ where $\mathfrak{i}_0$ is defined in (3.6). The criterion (3.9) becomes $\mathbf{b}V = 0$ for all $\mathbf{b} \in \mathbf{i}_0$ and $\mathbf{a}V \gg 0$. The level sets of such a V are the zero-time orbits $\mathbf{H}\xi$ and are trap boundaries if the trajectory from ξ never returns to $\mathbf{H}\xi$. With this in mind, if $\mathcal{P}_0^+(x)$ is nontrivial then it may be a candidate gauge polynomial.

Example 3.2. The control system of Example 1.7 is isomorphic to the scalar system $\dot{z} = (1 + ui)z$ on $\mathbb{C}_*$, whose Lie algebra is $\mathbb{C}$; the Lie group $\mathbf{G}_\alpha = \mathbb{C}_*$ is transitive on $\mathbb{C}_*^2$ and the zero-time ideal is $\mathfrak{i}_0 = \mathbb{R}i$. Then $V(z) := |z|^2$ is a good gauge function: $\dot{V} = 2V \gg 0$. For every $\rho > 0$ the set $\{z | V(z) \geq \rho\}$ is a trap. △

Example 3.3. Let $\mathfrak{g} = \{A, B\}_\mathcal{L}$ and suppose that either (a) $\mathfrak{g} \simeq \mathfrak{so}(n) \oplus \mathbb{R}$ or (b) $\mathfrak{g} \simeq \mathfrak{so}(p, q) \oplus \mathbb{R}$. Both types of Lie algebra have the following properties, that are easily verified.

There exists a nonsingular symmetric matrix Q such that for each $X \in \mathfrak{g}$ there exists a real eigenvalue λ_X such that $X^\top Q + QX = \lambda_X Q$ is satisfied, and $\mathcal{P}_0^+(x) = x^\top Qx$. In case (a) $Q \gg 0$; then $\mathfrak{g}$ is transitive (case I.1 in Section D.2). For case (b), $\mathfrak{g} = \mathfrak{so}(p, q) \oplus \mathbb{R}$ is weakly transitive.

Now suppose in either case that A and B are in $\mathfrak{g}$, $\lambda_A > 0$ and $\lambda_B = 0$. Let $V(x) := (x^\top Qx)^2$; then $V \gg 0$, $\mathbf{a}V \gg 0$, and $\mathbf{b}V = 0$, so there exists a trap for (3.5). In case (a) $\mathfrak{i}_0 = \mathfrak{so}(n)$ and in case (b) $\mathfrak{i}_0 = \mathfrak{so}(p, q)$. △

Exercise 3.2. This system has Lie rank 2 but has a trap; find it with a quadratic gauge function $V(x)$.

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x + u \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} x \qquad \triangle$$

Example 3.4. The positive orthant $\mathbb{R}_+^n := \{x | x_i > 0,\ i \in 1, \dots, n\}$ is a trap for $\dot{x} = Ax + uBx$ corresponding to the Lyapunov function $V(x) = x_1 \cdots x_n$ if for all $\mu \in \Omega$ $(\mathbf{a} + \mu\mathbf{b})V(x) > 0$ whenever $V(x) \neq 0$; that is, the entries of A and B satisfy (see Chapter 6 for the details) $a_{i,j} > 0$ and $b_{i,j} = 0$ whenever $i \neq j$. For a generalization to the invariance of the other orthants, using a graph-theoretic criterion on A, see Sachkov [229]. △

Other positivity properties of matrix semigroups can be used to find traps; the rest of this subsection leads in an elementary way to Example 3.5, which do Rocio et al. [75] described using some nontrivial semigroup theory.

Definition 3.3. For $C \in \mathbb{R}^{n \times n}$ and $d \in \mathbb{N}$ let $\mathbf{P}_d(C)$, denote the property that all of the d–minors of matrix C are positive. △

Lemma 3.2. *Given any two $n \times n$ matrices M, N, each d–minor of MN is the sum of $\binom{n}{d}$ products of d–minors of M and N;*

$$\boldsymbol{P}_d(M) \text{ and } \boldsymbol{P}_d(N) \implies \boldsymbol{P}_d(MN). \tag{3.10}$$

Proof. The first conclusion is a consequence of the Cauchy–Binet formula, and can be found in more precise form in Gantmacher [101, V. I §2]; or see Exercise 3.10. (The case $d = 1$ is just the rule for multiplying two matrices, row-into-column.) From that (3.10) follows. □

A semigroup whose matrices satisfy $\mathbf{P}_d$ is called here a $\mathbf{P}_d$ semigroup. Two examples are well known: $GL^+(n, \mathbb{R})$ is a $\mathbf{P}_n$ semigroup and is transitive on $\mathbb{R}^n_*$; and the $\mathbf{P}_1$ semigroup $\mathbb{R}^{n\times n}_+$ is described in Example 3.4. Let $\mathbf{S}[d, n]$ denote the largest semigroup in $\mathbb{R}^{n\times n}$ with property $\mathbf{P}_d$. For $d < n$, $\mathbf{S}[d, n]$ does not have an identity, nor does it contain any units, but I belongs to its closure; $\mathbf{S}[d, n] \bigcup \{I\}$ is a monoid.

Proposition 3.6. *If* $n > 2$, *then* $\mathbf{S}[2, n] \bigcup \{I\}$ *is not transitive on* $\mathbb{R}^n_*$.

Proof. It suffices to prove that $\mathbf{S}[2, 3] \bigcup \{I\}$ *is not transitive on* $\mathbb{R}^3_*$.

The proof is indirect. Assume that the semigroup $\mathbf{S} := \mathbf{S}[d, 3] \bigcup \{I\}$ is transitive on $\mathbb{R}^3_*$. Let the initial state be $\delta_1 := \mathrm{col}(1, 0, 0)$ and suppose that there exists $G \in \mathbf{S}$ such that $G\delta_1 = \mathrm{col}(p_1, -p_2, p_3)$ where the p_i are strictly positive numbers. The ordered signs of this vector are $\{+, -, +\}$. From the two leftmost columns of G form the three 2–minors $m_{i,j}$ from rows i, j, supposedly positive. The inequality $m_{i,j} > 0$ remains true after multiplying its terms by p_k, $k \notin \{i, j\}$:

$$G = \begin{bmatrix} p_1 & g_{1,2} & g_{1,3} \\ -p_2 & g_{2,2} & g_{2,3} \\ p_3 & g_{3,2} & g_{3,3} \end{bmatrix};$$

$$p_3 m_{1,2} = p_1 p_3 g_{2,2} + p_2 p_3 g_{1,2} > 0,$$
$$p_2 m_{1,3} = p_1 p_2 g_{2,3} - p_2 p_3 g_{1,2} > 0,$$
$$p_1 m_{2,3} = -(p_1 p_2 g_{3,2} + p_1 p_3 g_{1,2}) > 0.$$

Adding the first two inequalities, we obtain a contradiction of the third one. That shows that no element of the semigroup takes δ_1 to the open octants $\{x |\, x_1 > 0, x_2 < 0, x_3 > 0\}$ and $\{x |\, x_1 < 0, x_2 > 0, x_3 < 0\}$ that have, respectively, the sign orders $\{+, -, +\}$ and $\{-, +, -\}$.

That is enough to see what happens for larger n. The sign orders of inaccessible targets $G\delta_1$ have more than one sign change. That is, in the ordered set $\{g_{1,1}, g_{2,1}, g_{3,1}, g_{4,1}, \dots\}$, only one sign change can occur in $G\delta_1$; otherwise, there is a sequence of coordinates with one of the patterns $\{+, -, +\}$ or $\{-, +, -\}$ that leads to a contradiction.[11] □

Remark 3.1. **i.** For $n = 3$, the sign orders with at most one *sign change* are: $\{+, +, +\}$, $\{+, +, -\}$, $\{+, -, -\}$, $\{-, -, -\}$, $\{-, -, +\}$, $\{-, +, +\}$. These are consistent with $\mathbf{P}_2(G)$. That can be seen from these matrices in $\mathbf{S}[2, 3]$ and their negatives:

[11] Thanks to Luiz San Martin for pointing out the sign-change rule for the matrices in a $\mathbf{P}_2$ semigroup.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix}.$$

The marginal cases $\{0,+,0\}$, $\{+,0,0\}$, $\{0,0,+\}$, $\{+,0,+\}$, $\{0,+,+\}$, $\{+,+,0\}$ and their negatives are easily seen to give contradictions to the positivity of the minors.

ii. $\mathbf{S}[2,n]$ contains singular matrices. This matrix A defined by $a_{i,j} = j - i$ satisfies $\mathbf{P}_2$ because for positive i, j, p, q the 2-minor $a_{i,j}a_{i+p,j+q} - a_{i,j+q}a\sqrt{i+pj} = pq$ and $pq > 0$; but since the ith row is drawn from an arithmetic sequence $j + \phi(i)$, the matrix A kills any vector like $\mathrm{col}(1,-2,1)$ or $\mathrm{col}(1,-1,0,-1,1)$ that contains the constant difference and its negative. (For $\mathbf{P}_2$, it suffices that ϕ be an increasing function.) The semigroup $\mathbf{P}_2$ matrices with determinant 1 is $\mathbf{S}[2,n] \cap \mathrm{SL}(n,\mathbb{R})$.

The matrices that are $\mathbf{P}_d$ for all $d \le n$ are called totally positive. An entertaining family of totally positive matrices with unit determinant has antidiagonals that are partial rows of the Pascal triangle:

$$a_{i,j} = \frac{(j+i-2)!}{(j-1)!(i-1)!}, \quad i,j \in 1,\ldots,n; \quad \text{for example} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}. \qquad \triangle$$

Proposition 3.7. *Let* $\mathbf{S}$ *be the matrix semigroup generated by*

$$\dot{X} = AX + uBX, \ u \in \mathcal{PK}, X(0) = I. \tag{3.11}$$

If for each $\mu \in \Omega$ *there exists* $t > 0$ *such that* $I + t(A + \mu B)$ *satisfies* $\mathbf{P}_d$ *then* $\mathbf{P}_d(X)$ *for each matrix* X *in* $\mathbf{S}\setminus\{I\}$.

Proof. For any matrix C such that $\mathbf{P}_d(I + tC)$ there exists $\tau > 0$ such that $\mathbf{P}_d(\exp(\tau C))$. Applying (3.10) $\mathbf{P}_d(\exp(t(A + \mu B)))$ for all $t \ge 0$ and $\mu \in \Omega$; therefore, each matrix in $\mathbf{S}$ except I satisfies $\mathbf{P}_d$. □

The following is Proposition 2.1 of do Rocio et al. [75], from which Example 3.5 is also taken.

Proposition 3.8. *Let* $\mathbf{S} \subset \mathrm{SL}(n,\mathbb{R})$ *be a semigroup with nonempty interior. Then* $\mathbf{S}$ *is transitive on* $\mathbb{R}^n_*$ *if and only if* $\mathbf{S} = \mathrm{SL}(n,\mathbb{R})$.

Definition 3.4 ([148]). The matrix $B \in \mathfrak{gl}(n,\mathbb{R})$ is strongly regular if its eigenvalues $\lambda_k = \alpha_k + i\beta_k$, $k \in 1,\ldots,n$ are distinct, including $2m$ conjugate-complex pairs, and the real parts $\alpha_1 < \ldots < \alpha_{n-m}$ satisfy $\alpha_i - \alpha_j \ne \alpha_p - \alpha_q$ unless $i = p$ and $j = q$. △

The strongly regular matrices are an open dense subset of $\mathbb{R}^{n\times n}$: the conditions in Definition 3.4 hold unless certain polynomials in the λ_i vanish. (The number $\alpha_i = \Re(\lambda_i)$ is either an eigenvalue or the half the sum of two eigenvalues, since B is real.)

Example 3.5. Using `LieTree`, you can easily verify that for the system $\dot{x} = Ax + uBx$ on $\mathbb{R}^4_*$ with

$$A = \begin{bmatrix} 0 & 1 & 0 & -2 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -2 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}, \quad \{A, B\}_{\mathcal{L}} = \mathfrak{sl}(4, \mathbb{R}),$$

so the interior of $\mathbf{S}$ is not empty. For any μ, for sufficiently small t the 2–minors of $I + t(A + \mu B)$ are all positive. From Proposition 3.7, $\mathbf{P}_2$ holds for each matrix in $\mathbf{S}$. Thus $\mathbf{S}$ is a proper semigroup of $\mathrm{SL}(4, \mathbb{R})$. It has no invariant cones in $\mathbb{R}^4_*$. That it is not transitive on $\mathbb{R}^4_*$ is shown by Proposition 3.6. The trap U lies in the complement of the set of orthants whose sign orders have more than one sign change, as shown in Proposition 3.6.

With the given A, the needed properties of $\mathbf{S}$ can be obtained using any $B = \mathrm{diag}(b_1, \dots, b_4)$ such that $\mathrm{tr}(B) = 0$ and B is *strongly regular*. Compare the proof of Theorem 2.3.

△

3.3.3 *Transitive Semigroups*

If $\mathbf{S}_\Omega = \mathbf{G}_\alpha$, evidently the identity is in its interior $\overset{\circ}{\mathbf{S}}_\Omega$. The following converse is useful because if $\mathbf{G}_\alpha$ is one of the groups transitive on $\mathbb{R}^n_*$ (Appendix D). then $\mathbf{S}$ for (3.2) is transitive on $\mathbb{R}^n_*$.

Theorem 3.2. *If the identity is in* $\overset{\circ}{\mathbf{S}}_\Omega$ *then* $\mathbf{S}_\Omega = \mathbf{G}_\alpha$.

Proof. By its definition as a set of transition matrices, $\mathbf{G}_\alpha$ is a connected set; if we show that $\mathbf{S}_\Omega$ is both open and closed in $\mathbf{G}_\alpha$, then $\mathbf{S}_\Omega = \mathbf{G}_\alpha$ will follow.

Suppose $I \in \overset{\circ}{\mathbf{S}}_\Omega$. Then there is a neighborhood $\mathbf{U}$ of I in $\mathbf{S}_\Omega$. Any $X \in \mathbf{S}_\Omega$ has a neighborhood $XU \subset \mathbf{S}_\Omega$; therefore, $\mathbf{S}_\Omega$ is open.

Let $Y \in \partial\mathbf{S}_\Omega$ and let $\{Y_i \in \mathbf{S}_\Omega\}$ be a sequence of matrices such that $Y_i \to Y$; by Proposition 1.5, the Y_i are nonsingular. Then $YY_i^{-1} \to I$, so for some j we have $YY_j^{-1} \in U \subset \mathbf{S}_\Omega$. By the semigroup property $Y = (YY_j^{-1})Y_j \in \mathbf{S}_\Omega$, so $\mathbf{S}_\Omega$ is closed in $\mathbf{G}_\alpha$. □

The following result is classical in control theory; see Lobry [189] or Jurdjevic [147, Ch. 6]. The essential step in its proof is to show that I_n is in the closure of $\mathbf{S}_\Omega$.

Proposition 3.9. *If the matrix Lie group* $\mathbf{K}$ *is compact with Lie algebra* $\mathfrak{k}$ *and the matrices in* (3.2) *satisfy* $\{A, u_1B_1, \dots, u_mB_m | \, u \in \Omega\}_{\mathcal{L}} = \mathfrak{k}$ *then* $\mathbf{S}_\Omega = \mathbf{K}$.

Example 3.6. The system $\dot{z} = (i + 1)z + uz$ on $\mathbb{C}_*$ generates the Lie algebra $\mathbb{C}$; $\mathbf{G}_\alpha = \mathbb{C}_*$ and $\mathfrak{i}_0 = \mathbb{R}$. If $\Omega = \mathbb{R}$ the system is controllable. If $\Omega := [-\mu, \mu]$ for

$0 < \mu \leq 1$ then controllability fails. The semigroup $\mathbf{S}_\Omega$ is $\{1\} \bigcup \{z \,|\, |z| > 1\}$, so 1 is not in its interior. If $\mu = 1$, the unit circle in $\mathbb{C}_*$ is the boundary of $\mathbf{S}_\Omega$. △

Definition 3.5. A nonzero matrix $C \in \mathbb{R}^{n\times n}$ will be called neutral[12] if $C^\top Q + QC = 0$ for some $Q \gg 0$.

Cheng [57] pointed out the importance of neutral matrices in the study of controllability of bilinear systems; their properties are listed in the following lemma, whose proof is an easy exercise in matrix theory.

Lemma 3.3 (Neutral matrices). *Neutrality of C is equivalent to each of the following properties:*

1. *C is similar over $\mathbb{R}$ to a skew-symmetric matrix.*
2. spec(C) *lies in the imaginary axis and C is diagonalizable over $\mathbb{C}$.*
3. *The closure of $e^{\mathbb{R}C}$ is compact.*
4. *There exists $\sigma > 0$ and a sequence of times $\{t_k\}_1^\infty$ with $t_k \geq \sigma$ such that* $\lim_{k\to\infty} \|e^{t_k C} - I\| = 0$.
5. $|\operatorname{tr}(C)| = 0$ *and* $\|e^{tC}\|$ *are bounded on* $\mathbb{R}$.

Theorem 3.3. *Let $\Omega = \mathbb{R}^m$.*
(i) If there are constants $\delta_i \in \Omega$ such that $A_\delta := A + \sum_1^m \delta_i B_i$ is neutral then $\mathbf{S} = \mathbf{G}_\alpha$*;*
(ii) If, also, the Lie rank of $\mathfrak{g}_\alpha$ is n then (3.2) *is controllable.*

Proof. [13] If (i) is proved then (ii) follows immediately. For (i) refer to Proposition 2.7, rewriting (2.21) correspondingly, and apply it to

$$\dot{X} = u_0 A_\delta X + u_1 B_1 X;$$

the generalization for $m > 1$ is straightforward.

To return to (3.2) impose the constraint $u_0 = 1$, and a problem arises. We need first to obtain an approximation $C \in \mathfrak{g}_\alpha$ to $[A_\delta, B_1]$ that satisfies, for some τ in an interval $[0, \tau_1]$ depending on A_δ and B_1,

$$C = e^{\tau A_\delta} B_1 e^{-\tau A_\delta}.$$

If $e^{\tau A_\delta}$ has commensurable eigenvalues, it has some period T; for any $\tau > 0$ and $k \in \mathbb{N}$ such that $kT > \tau$, $e^{-\tau A_\delta} = e^{(kT-\tau)A_\delta} \in \mathbf{S}$. In the general (dense) case, there are open intervals of times $t > \tau$ such that $e^{(t-\tau)A_\delta} \in \mathbf{S}$. For longer products like (2.21), this approximation must be made for each of the $\exp(-\tau A)$ factors; the $\exp(-\tau u B)$ factors raise no problem. □

---

[12] Calling C *neutral*, meaning "similar to skew-symmetric," refers to the neutral stability of $\dot{x} = Cx$ and is succinct.

[13] The proof of Theorem 3.3 in Cheng [56] is an application of Jurdjevic and Sussmann [151, Th. 6.6].

Proposition 3.10.[14] *Suppose $\dot{X} = AX + uBX$, for $n = 2$, has characteristic polynomial $\boldsymbol{p}_{A+\mu B}(s) := \det(sI - (A + uB)) = s^2 - \alpha s + \beta_0 + \beta_1 u$ with $\beta_1 \neq 0$. For any α, the semigroup $\mathbf{S}(\alpha) = \{X(t;u) |\, t \in \mathbb{R}, u(t) \in \mathbb{R}\}$ is transitive on $\mathbb{R}^2_*$; but for $\alpha \neq 0$, $\mathbf{S}(\alpha)$ is not a group.*

Proof. Choose a control value μ_1 such that $\beta_0 + \beta_1\mu_1 < 0$; then the eigenvalues of $A + \mu_1 B$ are real and of opposite sign. Choose μ_2 such that $\mu_2 > |\alpha|/2$; the eigenvalues are complex. The corresponding orbits of $\dot{x} = (A + \mu B)x$ are hyperbolas and (growing or shrinking) spirals. A trip from any ξ to any state $\bar{x}$ is always possible in three steps. From $x(0) = \xi$, choose $u(0) = \mu_2$ and keep that control until the spiral's intersection at $t = t_1$ with one of the four eigenvector rays for μ_1; for $t > t_1$ use $u(t) = \mu_1$, increasing or decreasing $\|x(t)\|$ according to which ray is chosen; and choose $t_2 = t_1 + \tau$ such that $\exp(\tau(A + \mu_1 B)x(t_1) = \exp(-\tau(A + \mu_2)\bar{x}$. See Example 3.7, where $\bar{x} = \xi$, and Fig. 3.2.

Since $\operatorname{tr}(A + uB) = \alpha \neq 0$, $\det(X(t;u)) = e^{t\alpha}$ has only one sign. Therefore, $\mathbf{S}(\alpha)$ is not a group for $\alpha \neq 0$; I is on its boundary. □

Remark 3.2. Thanks to L. San Martin (private communication) for the following observation. In the system of Proposition 3.10, if $\alpha = 0$ then $\mathbf{S}(0) = SL(2,\mathbb{R})$ (which follows from Braga Barros et al. [23] plus these observations: $\overset{\circ}{\mathbf{S}}(0)$ is nonempty and transitive on $\mathbb{R}^2_*$); also

$$GL^+(2,\mathbb{R}) = SL(2,\mathbb{R}) \bigcup \{\mathbf{S}(\alpha), \alpha < 0\} \bigcup \{\mathbf{S}(\alpha), \alpha > 0\}. \qquad \triangle$$

Example 3.7. Let

$$\dot{x} = \begin{bmatrix} -1 & 1 \\ u & 0 \end{bmatrix} x; \ \boldsymbol{p}(s) = s^2 + s - u. \tag{3.12}$$

This system is controllable on $\mathbb{R}^2_*$ using the two constant controls $\mu_1 = 2$, $\mu_2 = -3$. The trajectories are stable spirals when $u = \mu_2$. For $u = \mu_1$, they are hyperbolas; the stable eigenvalue -2 has eigenvector col$(-1, 1)$ and the unstable eigenvalue 2 has eigenvector col$(1, 2)$). Sample orbits are shown in Fig. 3.2. As an exercise, using the figure, find at least two trajectories leading from ξ to ξ.

Notes for this example: (i) any negative constant control is a stabilizer for (3.12); (ii) the construction given here is possible if and only if $-1/4 \in \overset{\circ}{\Omega}$; and (iii) do Exercise 3.4. △

Remark 3.3. The time-reversal of system (3.1) is

$$\dot{x} = -Ax + \sum_{i=1}^{m} v_i(t) B_i x, \ v(t) \in -\Omega. \tag{3.13}$$

[14] Proposition 3.10 has not previously appeared.

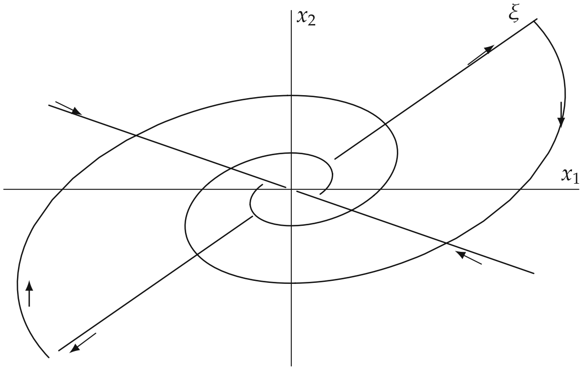


Fig. 3.2. Example 3.7: trajectories along the unstable and stable eigenvectors for μ_1, and two of the stable spiral trajectories for μ_2. (Vertical scale = 1/2 horizontal scale.)

If using (3.1) one has $\zeta = X(T;u)\xi$ for $T > 0$ then ξ can be reached from ζ by the system (3.13) with $v_i(t) = -u_i(T-t)$, $0 \le t \le T$; so the time-reversed system is controllable on $\mathbb{R}^n_*$ if and only if the system (3.1) is controllable. △

Exercise 3.3. This system is analogous to Example 3.7. It lives on $\mathbb{R}^3_*$ and has three controls, one of which is constrained:

$$\dot{x} = F(u,v,w)x,\ F := \begin{bmatrix} u & -1 & v \\ 1 & u & w \\ -v & -w & -3 \end{bmatrix} x,\ |u| < 1,\ (v,w) \in \mathbb{R}^2.$$

Since $\operatorname{tr} F(u,v,w) = 2u - 3 < 0$, for $t > 0$ the transition matrices satisfy $\det(X(t;u)) < 1$. Therefore $\mathbf{S}$ is not a group. Verify that the Lie algebra $\{F(u,v,w)\}_{\mathcal{L}}$ is transitive and show geometrically[15] that this system is controllable. △

3.3.3.1 Hypersurface Systems

If $m = n-1$ and $A, B_1, \dots, B_{n-1}$ are linearly independent, then (3.1) is called a hypersurface system[16] and its attainable sets are easily studied. The most understandable results use unbounded controls; informally, as Hunt [137] points out, one can use impulsive controls (Schwartz distributions) to achieve transitivity on each $n-1$-dimensional orbit $\mathbf{G}_\beta\xi$. Suppose that each such orbit is a hypersurface H that separates $\mathbb{R}^n_*$ into interior and exterior regions. It is then sufficient for controllability that trajectories with $u = 0$ enter and leave

[15] Hint for Exercise 3.3: The radius can be changed with u, for trajectories near the (x_1, x_2)-plane; the angular position can be changed using the drift term and large values of v, w. Problem: Is there a similar example on $\mathbb{R}^3_*$ using only two controls?

[16] See the work of Hunt [136, 137] on hypersurface systems on real-analytic and smooth systems. The results become simpler for bilinear systems.

these regions. For bilinear systems, one would like $\mathbf{G}_\beta$ to be compact so that H is an ellipsoid.

Example 3.8. Suppose $n = 2$ and $\Omega = \mathbb{R}$. If A has real eigenvalues $\alpha_1 > 0, \alpha_2 < 0$ and B has eigenvalues $\pm\beta i$ then (3.5) is controllable on $\mathbb{R}^2_*$.

Since $\det(sI - A - uB) = s^2 + (\alpha_2 - \alpha_1)s + \beta^2 u^2$. For $u = 0$, there are real distinct eigenvectors ξ, ζ such that $A\xi = \alpha_1\xi$ and $A\zeta = \alpha_2\zeta$. These two directions permit arbitrary increase or decrease in $x^\top x$ with no change in the polar angle θ. If u is sufficiently large, the system can be approximated by $\dot{x} = uBx$; θ can be arbitrarily changed without much change in $x^\top x$. △

Problem 3.1. To generalize Example 3.8 to $n > 2$, show that system (3.1) is controllable on $\mathbb{R}^n_*$ if $\Omega = \mathbb{R}^m$, $A = \operatorname{diag}(\alpha_1, \dots \alpha_n)$ has both positive and negative eigenvalues, and $\mathbf{G}_\beta$ is one of the groups listed in Equation (D.2) that are transitive on spheres, such as $\mathrm{SO}(n)$. △

3.4 Accessibility

If the neutrality hypothesis (i) of Theorem 3.3 is not satisfied, the Lie rank condition on $\mathfrak{g}$ is not enough for accessibility,[17] much less strong accessibility for (3.1); so a stronger algebraic condition, the ad-condition, will be derived for single-input systems (3.5). (The multiple-input version can be found at the end of Section 3.4.1.) On the other hand, if the drift matrix A is neutral, Section 3.5 shows how to apply the ad-condition to control with small bounds. When we return to stabilization in Section 3.6, armed with the ad-condition we can find stabilizing controls that are smooth on $\mathbb{R}^n$ or at least smooth on $\mathbb{R}^n_*$.

3.4.1 Openness Conditions

For real-analytic control systems, it has been known since [80, 251] and independently [269] that controllability on a simply connected manifold implies strong accessibility. (Example 1.8 shows what happens on $\mathbb{R}^2_*$.) Thus, given a controllable bilinear system (3.5) with $n > 2$, for each state x and each $t > 0$ the attainable set $\mathcal{A}_{t,\Omega}(x)$ contains a neighborhood of $\exp(tA)x$. A way to algebraically verify strong accessibility will be useful.

We begin with Theorem 3.4; it shows the effect of a specially designed control perturbation $u_i(t)B_i$ in (3.1) on the endpoint of a trajectory of $\dot{x} = Ax$ while all the other inputs $u_j, j \neq i$ are zero. It suffices to do this for $B_1 = B$ in the single-input system (3.5).

Suppose that the minimum polynomial of ad_A (Section A.2.2) is

[17] See Example 1.7 and Definition 1.10.

$$P(\lambda) := \lambda^{\mathbf{r}} + a_{\mathbf{r}-1}\lambda^{\mathbf{r}-1} + \cdots + a_0;$$

its degree $\mathbf{r}$ is less than n^2.

Lemma 3.4. *There exist* $\mathbf{r}$ *real linearly independent exponential polynomials* f_i *such that*

$$e^{-t\,\mathrm{ad}_A} = \sum_{i=0}^{\boldsymbol{r}-1} f_i(t)\,\mathrm{ad}_A^i; \tag{3.14}$$

$$P(d/dt)f_i = 0 \ \ and \ \ \frac{d^j f_i(0)}{dt^j} = \delta_{i,j},\ i \in 0,\ldots,\boldsymbol{r}-1,\ j \in 0,\ldots,\boldsymbol{r}-1.$$

Proof. The lemma is a consequence of Proposition 1.2 and the existence theorem for linear differential equations, because $f^{(r)}(t) + a_{r-1}f^{(r-1)}(t) + \cdots + a_0 f(t) = 0$ can be written as a linear dynamical system $\dot{x} = Cx$ on $\mathbb{R}^r$ where $x_i := f^{(r-i)}$ and C is a companion matrix. □

Let $\mathfrak{F}$ denote the kernel in C^ω of the operator $P(\frac{d}{dt})$; the dimension of $\mathfrak{F}$ is $\mathbf{r}$ and $\{f_1,\ldots,f_r\}$ is a basis for it. Choose for $\mathfrak{F}$ a second basis $\{g_0,\ldots,g_{\mathbf{r}-1}\}$ that is dual over $[0,T]$; that is, $\int_0^T f_i(t)g_j(t)\,dt = \delta_{i,j}$. For any $\eta \in \mathbb{R}^{\mathbf{r}}$ define

$$u(t;\eta) := \sum_{j=0}^{\mathbf{r}-1} \eta_j g_j(t) \ \text{ and } \alpha_k(t;\eta) := \int_0^t u(s;\eta) f_k(s)\,ds; \tag{3.15}$$

$$\text{note that } \ \frac{\partial \alpha_k(T;\eta)}{\partial \eta_i} = \delta_{i,k}. \tag{3.16}$$

Lemma 3.5. *For* η, u, α *as in* (3.15)*, let* $X(t;\epsilon u)$ *denote the solution of*

$$\dot{X} = (A + \epsilon u(t;\eta)B)X,\ X(0) = I,\ 0 \le t \le T.\ \textit{Then}$$

$$\frac{d}{d\epsilon}X(t;\epsilon u)|_{\epsilon=0} = e^{tA}\sum_{k=0}^{r} \alpha_k(t;\eta)\,\mathrm{ad}_A^k(B). \tag{3.17}$$

Proof. First notice that the limit $\lim_{\epsilon\to 0} X(\epsilon u;t) = e^{tA}$ is uniform in t on $[0,T]$. To obtain the desired derivative with respect to ϵ one uses the fundamental theorem of calculus and the pullback formula $\exp(t\,\mathrm{ad}_A)(B) = \exp(tA)B\exp(-tA)$ of Proposition 2.2 to evaluate the limit

$$\frac{1}{\epsilon}(X(t;\epsilon u) - e^{tA}) = \frac{1}{\epsilon}\int_0^t \frac{d}{ds}\Big(e^{(t-s)A}X(s;\epsilon u)\Big)\,ds$$
$$= \frac{1}{\epsilon}\int_0^t e^{(t-s)A}\Big((A+\epsilon u(s)B)X(\epsilon u;s) - AX(\epsilon u;s)\Big)\,ds$$
$$= e^{tA}\int_0^t u(s)e^{-sA}BX(\epsilon u;s)\,ds \xrightarrow[\epsilon\to 0]{} e^{tA}\int_0^t e^{-s\,\mathrm{ad}_A}(B)u(s)\,ds.$$

From (3.14)

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}(X(t;\epsilon u) - e^{tA}) = e^{tA}\sum_{i=1}^{\mathbf{r}}\left(\mathrm{ad}_A^i(B)\int_0^t u(s)f_i(s)\,ds\right),$$

and using (3.15) the proof of (3.17) is complete. □

Definition 3.6.
(i) System (3.5) satisfies the ad-condition[18] if for all $x \in \mathbb{R}^n_*$

$$\operatorname{rank} M_a(x) = n \text{ where } M_a(x) := \Big[Bx \;\; \mathrm{ad}_A(B)x \;\cdots\; \mathrm{ad}_A^{\mathbf{r}}(B)x\Big]. \tag{3.18}$$

(ii) System (3.5) satisfies the weak ad-condition if for all $x \in \mathbb{R}^n_*$

$$\operatorname{rank} M_a^w(x) = n;\; M_a^w(x) := \Big[Ax \;\; Bx \;\; \mathrm{ad}_A(B)x \;\cdots\; \mathrm{ad}_A^{\mathbf{r}}(B)x\Big]. \tag{3.19}$$

△

The lists $\{B, \mathrm{ad}_A(B), \dots, \mathrm{ad}_A^{\mathbf{r}}(B)\}$, $\{A, B, \mathrm{ad}_A(B), \dots, \mathrm{ad}_A^{\mathbf{r}}(B)\}$ used in the two ad-conditions are the matrix images of the Lie monomials in (2.13) that are of degree one in B; their spans usually are not Lie algebras. Either of the ad-conditions (3.18) and (3.19) implies that the Lie algebra $\{A, B\}_{\mathcal{L}}$ satisfies the LARC, but the converses fail. The next theorem shows the use of the two ad-conditions; see Definition 1.10 for the two kinds of accessibility.

Theorem 3.4. *(**I**) The ad-condition* (3.18) *for* (3.5) *implies strong accessibility: for all $\xi \neq 0$ and all $t > 0$, the set $\mathcal{A}_{t,\Omega}(\xi)$ contains an open neighborhood of $e^{tA}\xi$.*
*(**II**) The weak ad-condition* (3.19) *is sufficient for accessibility: indeed, for all $\xi \neq 0$ and $T > 0$ the set $\mathcal{A}^T_\Omega(\xi)$ has open interior.*

Proof. To show (**I**) use an input $u := u(t;\eta)$ defined by (3.15). From (3.17), the definition (3.15) and relation (3.16)

$$\frac{\partial X(t;u)}{\partial \eta_i}\Big|_{\eta=0} = \mathrm{ad}_A^i(B),\; i \in 0,\dots,\mathbf{r}-1, \tag{3.20}$$

[18] The matrices $M_a(x), M_a^w(x)$ are analogous to the Lie rank matrix $M(x)$ of Section 2.4.2. The ad-condition had its origin in work on Pontryagin's Maximum Principle and the bang-bang control problem by various authors. The definition in (3.18) is from Cheng [56, 57], which acknowledges a similar idea of Krener [165]. The best-known statement of the ad-condition is in Jurdjevic and Quinn [149].

so at $\eta = 0$, the mapping $\mathbb{R}^{\mathbf{r}} \to X(t;u)\xi$ has rank n for each ξ and each $t > 0$. By the Implicit Function Theorem, a sufficiently small open ball in $\mathbb{R}^{\mathbf{r}}$ maps to a neighborhood $e^{tA}\xi$, so $\mathcal{A}_{t,\Omega}(\xi)$ contains an open ball $\mathcal{B}(e^{tA}\xi)$. Please note that (3.20) is linear in B and that $\mathcal{B}(e^{tA}\lambda\xi) = \lambda\mathcal{B}(e^{tA}\xi)$.

To prove (**II**), use the control perturbations of part (**I**) and also use a perturbation of the time along the trajectory; for $T > 0$

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(X(T;0) - X(T-\epsilon;0)\right) = A.$$

Using the weak ad-condition, the image under $X(t;u)$ of $\Omega \times [0,T]$ has open interior. □

The ad-condition guarantees strong accessibility (as well as the Lie rank condition); $X(T;0)\xi$ is in the interior of $\mathcal{A}_\Omega(\xi)$. That does not suffice for controllability, as the following example shows by the use of the trap method of Section 3.3.2.

Example 3.9. On $\mathbb{R}^2$, let $\dot{x} = (A + uB)x$ with

$$A = \begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix},\ B = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix};\ M_a(X) = \begin{bmatrix}x_1 & -2x_2 & 4x_1 + 2x_2\\ -x_2 & 2x_1 & 2x_1 - 4x_2\end{bmatrix}.$$

The ideal generated by the 2×2 minors of $M(x)$ is $\langle x_1^2 - x_2^2, x_1x_2\rangle$. Therefore, $M_a(x)$ has rank two on $\mathbb{R}^2_*$; the ad-condition holds. The trajectories of $\dot{x} = Bx$ leave $V(x) := x_1x_2$ invariant; $\dot{V} = \mathbf{a}V + u\mathbf{b}V = x_1^2 + x_1x_2 + x_2^2 \gg 0$, so it is uncontrollable. △

To check either the strong (or weak) ad-condition apply the methods of Sections 2.5.2 and C.2.2: find the Gröbner basis of the $n\times n$ minors of $M_a(X)$ (or $M_a^w(X)$) and its real radical ideal $\sqrt[R]{\mathcal{J}}$ (or $\sqrt[R]{\mathcal{J}}^+$).

Proposition 3.11.
(i) If $\sqrt[R]{\mathcal{J}} = \langle x_1, \dots, x_n\rangle$, *then* $M(x)$ *has full rank on* $\mathbb{R}^n_*$ (ad-*condition).*
(ii) If $\sqrt[R]{\mathcal{J}}^+ = \langle x_1, \dots, x_n\rangle$, *then* $M_a^w(X)$ *has full rank on* $\mathbb{R}^n_*$ *(weak ad-condition).*
△

For reference, here are the two ad-conditions for multi-input control systems (3.1). Table 3.1 provides an algorithm to obtain the needed matrices. The ad-condition: for all $x \neq 0$

$$\operatorname{rank}\{\operatorname{ad}_A^i(B_j)x \,|\, 0 \le i \le \mathbf{r},\ 1 \le j \le m\} = n. \tag{3.21}$$

The weak *ad*-condition: add Ax to the list in (3.21):

$$\operatorname{rank}\{Ax, B_1x, \dots, B_mx, \operatorname{ad}_A(B_1)x, \dots, \operatorname{ad}_A^{\mathbf{r}}(B_m)x\} = n. \tag{3.22}$$

```
 «LieTree.m
AdCon::Usage="AdCon[A,L]: A, a matrix;
L, a list of matrices; outlist: {L, [A,L],...}."
AdCon[A_, L_]:= Module[
{len= Length[L], mlist={},temp, outlist=L,dp,dd,i,j,r,s},
{dp = Dimensions[A][[1]];
dd = dp*dp; temp = Array[te, dd];
For[i = 1, i < len + 1, i++,
{mlist = outlist; te[1] = L[[i]];
For[j = 1, j < dd, j++,
{r = Dim[mlist]; te[j + 1] = A.te[j] - te[j].A;
mlist = Append[mlist, te[j + 1]]; s = Dim[mlist];
If[s > r, outlist = Append[outlist, te[j + 1]], Break[]];
Print[j] } ] } ]; outlist}]
```

Table 3.1. Generating the ad-condition matrices for m inputs.

Theorem 3.4 does not take full advantage of multiple controls (3.1). Only first-order terms in the perturbations are used in (3.21), so brackets of degree higher than one in the B_i do not appear. There may be a constant control $u = \delta \in \Omega$ such that the ad-condition is satisfied by the set $\{\bar{A}, B_1, \dots, B_m\}$ where $\bar{A} = A + \delta_1 B_1 + \cdots + \delta_m B_m$; this is one way of introducing higher–order brackets in the B_i. For more about higher-order variations and their applications to optimal control, see Krener [165] and Hermes [125].

Exercise 3.4. Show that the ad-condition is satisfied by the system (3.12) of Example 3.7. △

3.5 Small Controls

Conditions that are sufficient for system (3.5) to be controllable for arbitrarily small control bounds are close to being necessary; this was shown in Cheng [56, 57] from which this section is adapted. That precisely the same conditions permit the construction of stabilizing controls was independently shown in Jurdjevic and Quinn [149].

Definition 3.7. If for every $\mu > 0$, the system $\dot{x} = Ax + uBx$ with $|u(t)| \le \mu$ is controllable on $\mathbb{R}^n_*$, it is called a small-controllable system. △

3.5.1 *Necessary Conditions*

To obtain a necessary condition for small-controllability Lyapunov's direct method is an inviting route. A negative domain for $V(x) := x^\top Q x$ is a nonempty set $\{x|\, V(x) < 0\}$; a positive domain is likewise a set $\{x|\, V(x) > 0\}$. By the continuity and homogeneity of V, such a domain is an open cone.

Lemma 3.6. *If there exists a nonzero vector* $z \in \mathbb{C}^n$ *such that* $Az = \lambda_1 z$ *with* $\Re(\lambda_1) > 0$ *then there exists* $\gamma > 0$ *and a symmetric matrix* R *such that* $A^\top R + RA = \gamma R - I$ *and* $V(x) := x^\top Rx$ *has a negative domain.*

Proof. Choose a number $0 < \gamma < \Re(\lambda_1)$ such that γ is not the real part of any eigenvalue of A and let $C := A - \gamma I$. Then $\mathrm{spec}(C) \cap \mathrm{spec}(-C) = \emptyset$, so (see Proposition A.1) the Lyapunov equation $C^\top R + RC = -I$ has a solution R. For that solution,

$$A^\top R + RA - 2\gamma R + I = 0. \text{ Let } z = \xi + \eta\sqrt{-1}, \text{ then}$$
$$0 = z^*(A^\top R + RA)z - 2\gamma z^* Rz + z^* z = (\bar{\lambda}_1 + \lambda_1 - 2\gamma)z^* Rz + z^* z;$$
$$\bar{\lambda}_1 + \lambda_1 > 2\gamma, \text{ so } z^* Rz = \frac{-z^* z}{\bar{\lambda}_1 + \lambda_1 - 2\gamma} < 0. \text{ That is,}$$
$$\xi^\top R\xi + \eta^\top R\eta = -\frac{\xi^\top \xi + \eta^\top \eta}{\bar{\lambda}_1 + \lambda_1 - 2\gamma} < 0,$$

so at least one of the quantities $V(\xi)$ or $V(\eta)$ is negative, establishing that V has a negative domain. □

Proposition 3.12. *If* (3.5) *is small-controllable then* $\mathrm{spec}(A)$ *is imaginary.*

Proof. [19] First suppose that A has at least one eigenvalue with positive real part, that R and γ are as in the proof of Lemma 3.6, and that $V(x) := x^\top Rx$. Along trajectories of (3.5)

$$\dot{V} = \gamma V(x) + W(x), \text{ where } W(x) := -x^\top x + 2ux^\top Bx.$$

Impose the constraint $|u(\cdot)| \le \mu$; then $W(x) < -x^\top x + 2\mu\|B\|x^\top x$, and if μ is sufficiently small, W is negative definite. Then $\dot{V} \le \mu V(x)$ for all controls such that $|u| < \mu$; so $V(x(t)) \le V(x(0))\exp(\mu t)$. Choose an initial condition such that $V(\eta) < 0$. All trajectories $x(t)$ starting at η satisfy $V(x(t)) < 0$, so states x such that $V(x) \ge V(\eta)$ cannot be reached, contradicting the hypothesis of small-controllability. Therefore, no eigenvalue of A can have a positive real part.

Second, suppose A has at least one eigenvalue with negative real part. If for a controllable system (3.5) we are given any ξ, ζ and a control u such that $X(T; u)\xi = \zeta$, then the reversed-time system (compare (1.40))

$$\dot{Z} = -AZ + u^- BZ, \; Z(0) = I, \text{ where } u^-(t) := -u(T - t), \; 0 \le t \le T$$

has $Z(T; u^-)\zeta = \xi$. Therefore, the system $\dot{x} = -Ax + uBx$ is controllable with the same control bound $|u(\cdot)| \le \mu$. Now $-A$ has at least one eigenvalue λ_1 with positive real part. Lemma 3.6 applies to this part of the proof, so a contradiction occurs; therefore, $\mathrm{spec}(A)$ is a subset of the imaginary axis. □

---

[19] This argument follows Cheng [57, Th. 3.7] which treats more general systems $\dot{x} = Ax + g(u, x)$.

Exercise 3.5. If A has an eigenvalue with positive real part, show that there exists $\mu > 0$ and a symmetric matrix R such that $A^\top R + RA = \mu R - I$, and $V(x) := x^\top Rx$ has a negative domain. Use this fact instead of the time-reversal argument in the proof of Proposition 3.12.

3.5.2 *Sufficiency*

The invariance of a sphere under rotation has an obvious generalization that will be needed in the Proposition of this section. Analogous to a spherical ball, given $Q \gg 0$ a Q-ball of size δ around x is defined as

$$\mathcal{B}_{Q,\delta}(x) := \{z |\, (z-x)^\top Q(z-x) < \delta^2\}.$$

If $A^\top Q + QA = 0$ then $\exp(tA)^\top Q \exp(tA) = Q$, and the Q-ball $\mathcal{B}_{Q,\delta}(x)$ is mapped by $x \mapsto \exp(tA)x$ to $\mathcal{B}_{Q,\delta}(x)$; that is,

$$\exp(tA)\mathcal{B}_{Q,\delta}(x) = \mathcal{B}_{Q,\delta}(\exp(tA)x). \tag{3.23}$$

Proposition 3.13 (Cheng [56, 57]). *If* (3.5) *satisfies the ad-condition and for some* $Q \gg 0$, $A^\top Q + QA = 0$, *then* (3.5) *is small-controllable on* $\mathbb{R}^n_*$.

Proof. Define $\rho := \|\xi\|$. From neutrality and Lemma 3.3 (Part 4), there exists a sequence of times $\mathcal{T} := \{t_i\}_1^\infty$ such that

$$X(t_k; 0)\xi = e^{t_k A}\xi \in \mathcal{B}_{Q,\rho/k}(\xi).$$

Take $\Omega := [-\mu, \mu]$ for any $\mu > 0$. By Theorem 3.4, for some $\delta > 0$ depending on μ and any $\tau > 0$ there is a Q-ball of size $\delta\rho$ such that

$$\mathcal{B}_{Q,\delta\rho}(e^{\tau A}\xi) \subset \mathcal{A}_{\tau,\Omega}(\xi).$$

From (3.23) that relation remains true for any $\tau = t_k \in \mathcal{T}$ such that $1/k < \delta/2$. Therefore,

$$\mathcal{B}_{Q,\rho/k}(\xi) \subset \mathcal{A}_{\tau,\Omega}(\xi).$$

Controllability follows from the Continuation Lemma (Section 1.6.2). □

Corollary 3.1. *If* (3.5) *satisfies the ad-condition and for some constant control* δ *the matrix* $A_\delta := A + \delta B$ *is neutral then* (3.5) *is controllable under the constraint* $|u(t) - \delta| < \mu$ *for any* $\mu > 0$.

Remark 3.4. Proposition 3.13 can be strengthened as suggested in Cheng [57]: replace the ad-condition by the weak ad-condition. In the above proof, one must replace the Q-ball by a small segment of a tube domain around the zero-control trajectory; this segment, like the Q-ball, is invariant under the mapping $\exp(tA)$. △

Example 3.10. Proposition 3.12 is too weak, at least for $n = 2$. For example, let

$$A := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ B := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix};\ M_a(X) := \begin{bmatrix} Bx\ [A,B]x \end{bmatrix} = \begin{bmatrix} x_2 & x_1 \\ x_1 & -x_2 \end{bmatrix},$$

the spectrum of A lies on the imaginary axis, A cannot be diagonalized, the ad-condition is satisfied, and $\boldsymbol{p}_{A+\mu B}(s) = s^2 - \mu(1+\mu)$. Use constant controls with $|u| = \mu < 1/2$. If $u = -\mu$, the orbits described by the trajectories are ellipses; if $u = \mu$, they are hyperbolas. By the same geometric method used in Example 3.7, this system is controllable; that is true no matter how small μ may be. △

Problem 3.2. Does small-controllability imply that A is neutral? △

3.6 Stabilization by State-Dependent Inputs

For the systems $\dot{x} = Ax + uBx$ in Section 3.2 with $u = \mu$, the origin is the only equilibrium and its basin of stability $\mathbf{U}$ is $\mathbb{R}^n_*$. With state-dependent control laws, the basin of stability may be small and there may be unstable equilibria on its boundary. For $n = 2$, this system has nontrivial equilibria for values of u given by $\det(A + uB) = q_0 s^2 + 2q_1 s + q_2 = 0$. If $q_0 \neq 0$ and $q_1^2 - q_0 q_2 > 0$, there are two distinct real roots, say μ_1, μ_2. If the feedback is linear, $\phi(x) = c^\top x$, there are two corresponding nonzero equilibria ξ, η given by the linear algebraic equations

$$(A + \mu_1 B)\xi = 0,\ c^\top \xi = \mu_1,$$
$$(A + \mu_2 B)\eta = 0,\ c^\top \eta = \mu_1.$$

Stabilization methods for bilinear systems often have used quadratic Lyapunov functions. Some work has used the bilinear case of the following well-known theorem of Jurdjevic and Quinn [149, Th. 2].[20] Since it uses the hypotheses of Proposition 3.13, it will later be improved to a small-controllable version, Theorem 3.8.

Theorem 3.5. *If $\dot{x} = Ax + uBx$ satisfies the ad-condition* (3.18) *and A is neutral, then there exists a feedback control $u = \phi(x)$ such that the origin is the unique global attractor of the dynamical system $\dot{x} = Ax + \phi(x)Bx$.*

Proof. There exists $Q \gg 0$ such that $QA + A^\top Q = 0$. A linear change of coordinates makes $Q = I$, $A + A^\top = 0$ and preserves the ad-condition. In these new coordinates, let $V(x) := x^\top x$, which is radially unbounded. B is not proportional to A (by the ad-condition). If we choose the feedback control

[20] Jurdjevic and Quinn [149, Th. 2] used the ad-condition as a hypothesis toward the stabilization by feedback control of $\dot{x} = Ax + ug(x)$ where g is real-analytic.

$\phi(x) := -x^\top Bx$, we obtain a cubic dynamical system $\dot{x} = f(x)$ where $f(x) = Ax - (x^\top Bx)Bx$. Along its trajectories, the Lie derivative of V is nonincreasing: $\mathbf{f}V(x) = -(x^\top Bx)^2 \leq 0$ for all $x \neq 0$. If $B \gg 0$, the origin is the unique global attractor for $\mathbf{f}$; if not, $\mathbf{f}V(x) = 0$ on a nontrivial set $\mathbf{U} = \{x|\, x^\top Bx = 0\}$. The trajectories of $\dot{x} = f(x)$ will approach (or remain in) an $\mathbf{f}$-invariant set $\mathbf{U}_0 \subset \mathbf{U}$. On $\mathbf{U}$, however, $\dot{x} = Ax$ so $\mathbf{U}_0$ must be $\mathbf{a}$-invariant; that fact can be expressed, using $A^\top = -A$, as

$$0 = x(t)^\top Bx(t) = \xi^\top e^{-tA} Be^{tA}\xi = \xi^\top e^{t\,\mathrm{ad}_A}(B)\xi.$$

Now use the ad-condition: $\mathbf{U}_0 = \{0\}$. By LaSalle's invariance principle (Proposition 1.9), the origin is globally asymptotically stable. □

Definition 3.8. The system (3.1) will be called a JQ system if A is neutral and the ad-condition is satisfied. △

The ad-condition and the neutrality of A are invariant under linear coordinate transformations, so the JQ property is invariant too. Other similarity invariants of (3.5) are the real canonical form of B and the coefficients of the characteristic polynomial $\boldsymbol{p}_{A+\mu B}(s)$.

Remark 3.5. Using the quadratic feedback in the above theorem, $\|x(t)\|^2$ approaches 0 slowly. For example, with $A^\top = -A$ let $B = I$ and $V(x) = x^\top x$. Using the feedback $\phi(x) = -V(x)$, $\dot{V} = -V^2$; if $V(\xi) = \mu$ then $V(x(t)) = \mu/(1 + 2t\mu)$.

This slowness can be avoided by a change in the method of Theorem 3.5, as well as by a weaker hypothesis; see Theorem 3.8. △

3.6.1 A Question

If a bilinear system is controllable on $\mathbb{R}^n_*$, can it be stabilized with a feedback that is smooth (a) on $\mathbb{R}^n$? (b) on $\mathbb{R}^n_*$?

Since this question has a positive answer in case (a) for linear control systems, it is worth asking for bilinear systems. Example 3.11 will answer part (a) of our question negatively, using the following well-known lemma.

Lemma 3.7. *If for each real μ at least one eigenvalue of $A + \mu B$ lies in the right half-plane then there is no C^1 mapping $\phi : \mathbb{R}^n \to \mathbb{R}$ for which the closed-loop system $\dot{x} = Ax + \phi(x)Bx$ is asymptotically stable at* 0.

Proof. The conclusion follows from the fact that any smooth feedback must satisfy $\phi(x) = \phi(0) + \phi_*(0)x + o(\|x\|)$; $\dot{x} = Ax + \phi(x)Bx$ has the same instability behavior as $\dot{x} = (A + \phi(0)B)x$ at the origin. This observation can be found in [14, 52, 155] as an application of Lyapunov's principal of stability or instability in the first approximation (see [156]). □

Example 3.11. Let

$$A_\delta = \begin{bmatrix} \delta & -1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

for δ small and positive. As in Example 3.7, $\dot{x} = A_\delta x + uBx$ is controllable: the trajectories for $u = 1$ are expanding spirals and for $u = -2$ the origin is a saddle point. The polynomial $\det(sI - A_\delta - \mu B) = s^2 - \delta s + \mu + 1$ always has at least one root with positive real part; by Lemma 3.7, no stabilizer ϕ smooth on $\mathbb{R}^n$ exists. △

Part (b) of the question remains as an open problem; there may be a stabilizer smooth on $\mathbb{R}^n_*$. For Example 3.11, experimentally the rational homogeneous (see Section 3.6.3) control $\phi(x) = -(x^\top Bx)/(x^\top x)$ is a stabilizer for $\delta < \delta_0 := 0.2532$; the dynamical system $\dot{x} = A_\delta x + \phi(x)Bx$ has oval orbits at $\delta \doteq \delta_0$.

3.6.2 *Critical and JQ Systems*

3.6.2.1 Critical Systems

Bacciotti and Boieri [14] describes an interesting class (open in the space of 2×2 matrix pairs) of *critical* two-dimensional bilinear systems that have no constant stabilizer. A system is critical if and only if

(i) there is an eigenvalue with nonnegative real part for each matrix in the the set $\{A + \mu B, \mu \in \mathbb{R}\}$ and
(ii) for some real μ the real parts of the eigenvalues of $A + \mu B$ are nonpositive.

Example 3.12. This critical system

$$\dot{x} = Ax + uBx, \quad A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

has $\det(Bx, [A, B]x) = x_1^2 + x_2^2$ and neutral A, so it is a JQ system and is stabilizable by Theorem 3.5. Specifically, let $V(x) := (x_1^2 + x_2^2)/2$. Then $V \gg 0$ and the stabilizing control $u = x_2^2 - x_1^2$ results in $\dot{V} = -(x_1^2 - x_2^2)^2$. This system falls into case ***a*** of Theorem 3.6 below.

It is also of type (2.2a) in Bacciotti and Boieri [14], critical with $\mathrm{tr}(B) = 0$. From their earlier work [15] (using symbolic computations), there exist linear stabilizers $\phi(x) := px_1 + qx_2$ ($|q| > |p|, pq > 0$). The equilibria $((p+q)^{-1}, (p+q)^{-1})$ and $((q-p)^{-1}, (p-q)^{-1})$ are on the boundary of $\mathbf{U}$. △

3.6.2.2 Jurdjevic & Quinn Systems

Theorem 3.6 is a summary of Chabour et al. [52], Theorems 2.1.2, 2.2.2, and 2.3.2; the three cases are distinguished by the real canonical form of B.[21] The basic plan of attack is to find constants μ such that $(A + \mu B, B)$ has the JQ property.

Theorem 3.6. *There exists $\phi(x) := \mu + x^\top Px$ that makes the origin globally asymptotically stable for $\dot{x} = Ax + \phi(x)Bx$ if one of the conditions **a**–**c** holds:*

$$\begin{aligned}
&\boldsymbol{a} : B \textit{ is diagonalizable over } \mathbb{R},\ tr(A) = 0 = \mathrm{tr}(B),\ \textit{and}\ \mathbf{p}_1 > 0;\\
&\boldsymbol{b} : B \sim \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},\ tr(A) = 0 = \mathrm{tr}(B),\ \ \textit{and}\ \ \mathrm{tr}(AB) \neq 0;\\
&\boldsymbol{c} : A \sim \begin{bmatrix} 0 & b \\ b & 0 \end{bmatrix},\ B \sim \begin{bmatrix} 0 & -\mu \\ \mu & 0 \end{bmatrix}.
\end{aligned}$$

Proof (Sketch). All these cases satisfy $(A, B) \approx (\tilde{A}, B)$ where $\dot{x} = \tilde{A}x + uBX$ is a JQ system and Theorem 3.5 applies; but the details differ.

($\boldsymbol{a}$) We can find a basis for $\mathbb{R}^2$ in which

$$A = \begin{bmatrix} a & c \\ b & -a \end{bmatrix},\ B = \begin{bmatrix} \lambda_1 & 0 \\ 0 & -\lambda_1 \end{bmatrix}. \quad \text{Let } \tilde{A} = A - \frac{a}{\lambda_1}B = \begin{bmatrix} 0 & c \\ b & 0 \end{bmatrix}.$$

The hypothesis on $\mathbf{p}_1$ implies here that $\det(\tilde{A})\det(B) < 0$; then $bc < 0$ so $\det(Bx, [\tilde{A}, B]x) = 2\lambda_1(bx_1^2 - t_2^2)$ is definite (the ad-condition). The JQ property is thus true for $\dot{x} = \tilde{A}x + uBx$ with

$$V(x) = \frac{1}{2}(|b|x_1^2 + |c|x_2^2); \quad \tilde{\phi}(x) := -\lambda_1(|b|x_1^2 - |c|x_2^2)$$

is its stabilizing control, so $\phi(x) := \tilde{\phi}(x) - a/\lambda_1$ stabilizes $\dot{x} = Ax + uBx$.

($\boldsymbol{b}$) The calculations for this double-root case are omitted; the ad-condition is satisfied because

$$\mathrm{rank}(Bx, \mathrm{ad}_A(B)x, \mathrm{ad}_A^2(B)x) = 2;$$
$$\phi(x) = \frac{-1}{b}(a^2 + bc + 1) - b^2x_1x_2 + abx_2^2.$$

($\boldsymbol{c}$) Let $\tilde{A} = A + (2b/\mu)B$, then $V(x) := (3x_1^2 + x_2^2)/2 \gg 0$ is constant on trajectories of $\dot{x} = \tilde{A}x$. To verify the ad-condition, show that two of the three 2×2 minors of $(Bx, \mathrm{ad}_A(B)x, \mathrm{ad}_A^2(B)x)$ have the origin as their only common

[21] The weak ad-condition (3.19), which includes Ax in the list, was called the *Jurdjevic and Quinn condition* by Chabour et al. [52, Def. 1.1] and was called the ad-condition by [103] and others; but like Jurdjevic and Quinn [149], the proofs in these papers need the strong version.

zero. Since $x \mapsto 2\mu x_1 x_2$ stabilizes $\dot{x} = \tilde{A}x + uBx$, use $\phi(x) := 2b/\mu + 2\mu x_1 x_2$ to stabilize the given system. □

Theorem 2.1.3 in [52], which is a marginal case, provides a quadratic stabilizer when B is similar to a real diagonal matrix, $\mathbf{p}_1 > 0$, $\mathbf{p}_2 < 0$, $\mathbf{p}_3 = 0$ and $\det(B) \le 0$, citing Bacciotti and Boieri [14], but the stabilization is not global and better results can be obtained using the stabilizers in the next section.

3.6.3 *Stabilizers, Homogeneous of Degree Zero*

Recalling Definition 2.4, a function $\phi : \mathbb{R}^n \to \mathbb{R}$ is called homogeneous of degree zero if it satisfies $\phi(\alpha x) = \phi(x)$ for all real α. To be useful as a stabilizer, it must also be bounded on $\mathbb{R}^n_*$; such functions will be called $\mathcal{H}_0$ functions. Constant functions are a subclass $\mathcal{H}_0^0 \subset \mathcal{H}_0$. Other subclasses we will use are $\mathcal{H}_0^k$, the functions of the form $p(x)/q(x)$ where p and q are both homogeneous of degree k. Examples include $c^\top x/\sqrt{x^\top Q x}$, $Q \gg 0$, called $\mathcal{H}_0^1$, and $(x^\top R x)/(x^\top Q x)$, $Q \gg 0$, called $\mathcal{H}_0^2$. The functions in $\mathcal{H}_0^k$, $k > 0$, are discontinuous at $x = 0$ but smooth on $\mathbb{R}^n_*$; they are continuous on the unit sphere and constant on radii. They have explicit bounds: for example, if $Q = I$

$$|c^\top x|/\|x\| \le \|c\|; \quad \lambda_{\min}(R) \le x^\top R x/(x^\top x) \le \lambda_{\max}(R).$$

If they stabilize at all, the stabilization is global, since the dynamical system is invariant under the mapping $x \mapsto \rho x$.

One virtue of $\mathcal{H}_0^k$ in two dimensions is that such a function can be expressed in terms of $\sin^k(\theta)$ and $\cos^k(\theta)$ where $x = r\cos(\theta)$, $y = r\sin(\theta)$. Polar coordinates are useful in looking for $\mathcal{H}_0^1$ stabilizers for non-JQ systems; see Bacciotti and Boieri [15] and the following example.

Example 3.13. If $\operatorname{tr}(B) = 0 = \operatorname{tr}(A)$, no pair $(A + \mu B, B)$ is JQ. Let

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} x + u \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x. \quad \text{In polar coordinates}$$

$$\dot{r} = r\cos(2\theta), \;\; \dot{\theta} = -2\cos(\theta)\sin(\theta) - u.$$

$$\text{For } |c| > 2, \text{ let } \; u = c\cos(\theta) = c\frac{x_1}{\|x\|} \in \mathcal{H}_0^1.$$

Then $\dot{\theta} = -(2\sin(\theta) + c)\cos(\theta)$ has only two equilibria in $[-\pi, \pi]$ located at $\theta = \pm\pi/2$ (where $\dot{r} = -1$), and one of them is θ-stable. The origin is globally asymptotically stable but for trajectories starting near the θ-unstable direction $r(t)$ initially increases. Edit the first few lines of the script in Exercise 3.11 to plot the trajectories. △

The following theorem is paraphrased from Chabour et al. [52, Th. 2.1.4, 2.2.3, 2.3.2]. See page 87 for the $\mathbf{p}_i$ invariants.

Theorem 3.7. *Systems* (3.5) *on* $\mathbb{R}^2_*$ *can be stabilized by a homogeneous degree-zero feedback provided that one of* $\boldsymbol{a}$, $\mathbf{b}$, *or* $\mathbf{c}$ *is true:*

$(\boldsymbol{a})$ $\mathbf{p}_4 < 0$ *and at least one of these is true:*
(1) $\det(B) < 0$ *and* $\mathbf{p}_1 \le 0$;
(2) $\det(B) = 0$ *and* $\mathbf{p}_1 = 0$ *and* $\mathbf{p}_3 \le 0$;
(3) $\det(B) \le 0$ *and* $\mathbf{p}_1 > 0$ *and* $\mathbf{p}_2 < 0$ *and* $\mathbf{p}_3 < 0$.
$(\boldsymbol{b})$ $\operatorname{tr}(B) = 0$, $\operatorname{tr}(A) > 0$ *and* $\operatorname{tr}(AB) \neq 0$.
$(\boldsymbol{c})$ $\chi(AB, AB) > 0$, $\operatorname{tr}(B) = 0$ *and* $\operatorname{tr}(A) > 0$.

The proofs use polar coordinates and give the feedback laws ϕ as functions of θ; the coefficients α and β are chosen in [52] so that $\theta(t)$ can be found explicitly and $r(t) \to 0$. For cases $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$, respectively:

$$(\boldsymbol{a}) \quad (A, B) \approx \left(\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\right), \; \lambda_1 > 0, \; \lambda_2 \le 0,$$
$$\phi(\theta) = \alpha \sin(2\theta) + \beta \sin(2\theta).$$

$$(\boldsymbol{b}) \quad (A, B) \approx \left(\begin{bmatrix} a & -b \\ b & -a \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right), \; a > 0,$$
$$\phi(\theta) = \alpha b \sin(2\theta), \; \alpha < 0.$$
$$(\boldsymbol{c}) \quad (A, B) \approx \left(\begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \begin{bmatrix} 0 & -\mu \\ \mu & 0 \end{bmatrix}\right),$$
$$\phi(\theta) = \frac{1}{\mu}\left(-b\cos(2\theta) + \cos\left(\theta + \frac{3\pi}{4}\right)\right).$$

3.6.4 Homogeneous JQ Systems

In the n-dimensional case of $\dot{x} = Ax + uBx$, we can now improve Theorem 3.5. Its hypothesis (ii) is just the ad-condition stated in a formulation using some quadratic forms that arise naturally from our choice of feedback.

Theorem 3.8. *If (i)* $A^\top + A = 0$ *and (ii) there exists* $k \in \mathbb{N}$ *for which the real affine variety* $\mathbf{V}(x^\top Bx, x^\top[A, B]x, \dots, x^\top(\operatorname{ad}_A)^k(B)x)$ *is the origin, then the origin is globally asymptotically stabilizable for* $\dot{x} = Ax + uBX$ *with an* $\mathcal{H}^2_0$ *feedback whose bounds can be prescribed.*

Proof. Note that $A^\top B + BA = BA - AB$. Let $V(x) := \frac{1}{2}x^\top x$ and $f(x) := Ax - \mu\phi(x)Bx$ (for any $\mu > 0$) with corresponding vector field $\mathbf{f}$. Choose the feedback

$$\phi(x) := \frac{x^\top Bx}{x^\top x}. \text{ Then } \mathbf{f}V(x) = -\mu\frac{(x^\top Bx)^2}{x^\top x} \le 0.$$

Let $W(x) := x^\top Bx$ and $\mathbf{U} := \{x|\, W(x) = 0\}$. On the zero-set $\mathbf{U}$ of $\mathbf{f}V(x)$, $f(x) = Ax$, so $\mathbf{f} = \mathbf{a} := x^\top A^\top \partial/\partial x$. $W(x(t)) = 0$ on some interval of time if its successive time derivatives vanish at $t = 0$; these are

$$W(x) = x^\top Bx,\ \mathbf{a}W(x) = x^\top[B, A]x,\ \ldots,\ (\mathbf{a})^k W(x) = x^\top(-\operatorname{ad}_A)^k(B)x.$$

The only $\mathbf{a}$-invariant subset $\mathbf{U}_0 \subset \mathbf{U}$ is the intersection of the zero-sets of the quadratic forms $\{(\mathbf{a})^k W(x), 0 \le k < n^2\}$; from hypothesis (ii), $\mathbf{U}_0 = \{0\}$. The global asymptotic stability is now a consequence of Proposition 1.9, LaSalle's Invariance Principle. To complete the conclusion, we observe that the control is bounded by $\mu \min \operatorname{spec}(B)$ and $\mu \max \operatorname{spec}(B)$. □

The denominator $x^\top x$ can be replaced with any $x^\top Rx$, $R \gg 0$. The resulting behaviors differ in ways that might be of interest for control system design. If A is neutral with respect to $Q \gg 0$, Theorem 3.8 can be applied after a preliminary transformation $x = Py$ such that $P^\top P = Q$.

Problem 3.3. [22] If a homogeneous bilinear system is locally stabilizable, is it possible to find a feedback (perhaps undefined at the origin) which stabilizes it globally?

3.6.5 *Practical Stability and Quadratic Dynamics*

Given a bilinear system (3.1) with m linear feedbacks $u_i := c_{(i)}^\top x$, the closed loop dynamical system is

$$\dot{x} = Ax + \sum_{i=1}^{m} x^\top c_{(i)} B_i x.$$

We have seen that stabilization of the origin is difficult to achieve or prove using such linear feedbacks; the most obvious difficulty is that the closed-loop dynamics may have multiple equilibria, limit cycles, or other attractors. However, a cruder kind of stability has been defined for such circumstances. A feedback control $\phi(x)$ is said to provide practical stabilization at 0 if there exists a Q-ball $\mathcal{B}_{Q,\epsilon}(0)$ such that all trajectories of (3.3) eventually enter and remain in the Q-ball. If we use a scaled feedback $\phi(\rho x)$, by the homogeneity of (3.5) the new Q-ball will be $\mathcal{B}_{Q,\epsilon/\rho}(0)$. That fact (Lemma 3 in [49]) has the interesting corollary that practical stability for a feedback that is homogeneous of degree zero implies global asymptotic stability.

[22] J.C. Vivalda, private communication, January 2006.

Čelikovský [49] showed practical stabilizability of rotated semistable second-order systems — those that satisfy $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$ in Example 3.1 — and some third-order systems.

Theorem 3.9 (Čelikovský [49], Th. 1). *For any system $\dot{x} = Ax + uBx$ on $\mathbb{R}^2$ that satisfies $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$ there exists a constant μ and a family of functions $\{c^\top x\}$ such that with $u := \mu + c^\top x$ every trajectory eventually enters a ball whose radius is of the order $o(\|c\|^{-1})$.*

Proof. For such systems, first one shows [49, Lemma 1] that there exists $\mu_0 \in \mathbb{R}$ and orthogonal matrix P such that

$$\hat{A} := P(A + \mu_0 B)P^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \lambda_1 > 0, \lambda_2 < 0;\ B = \beta \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

For $\dot{x} = (\hat{A} + uB)x$ let $u := (\gamma, 0)^\top x/\beta$, for any real γ, obtaining

$$\dot{x}_1 = \lambda x_1 - \gamma x_1 x_2, \quad \dot{x}_2 = \lambda_2 x_2 + \gamma x_1^2.$$

The gauge function $V(x) := \frac{1}{2}(x_1^2 + (x_2 - 2\lambda_1/\gamma)^2)$ is a Lyapunov function for this system:

$$\dot{V}(x) = -\lambda_1 x_1^2 + \lambda_2 \left((x_2 - \lambda_1/\gamma)^2 - (\lambda_1/\gamma)^2\right).$$

□

Theorem 3.10 (Čelikovský [49], Th. 3). *Using the notation of Theorem 3.9, consider bilinear systems* (3.5) *on* $\mathbb{R}^3_*$ *with* $\Omega = \mathbb{R}$ *and* $A = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_2)$, $B \in \mathrm{SO}(3)$, *where* $\lambda_1 > 0$, $\lambda_2 < 0$, $\lambda_3 < 0$. *If* $\gamma < 0$, *this system is globally practically stabilizable by a linear feedback.*

Proof. Use the feedback $u = x_1$, so we have the dynamical system $\dot{x} = Ax + x_1 Bx$ with A and B, respectively, the diagonal and skew-symmetric matrices in Example 3.1. The gauge function will be, with $k > 0$, the positive definite quadratic form

$$V(x) := \frac{1}{2}(x_1^2 + (x_2 - 2k\lambda_3 c_3)^2 + (x_3 + 2k\lambda_2 c_2)^2).$$

The Lie derivative $W(x) := \dot{V}(x)$ is easily evaluated, using $\gamma < 0$, to be a negative definite quadratic form plus a linear term, so there exists a translation $y = x - \xi$ that makes $W(y)$ a negative definite quadratic form. In the original coordinates, for large enough x there exists E, a level-ellipsoid of W that includes a neighborhood of the origin and into which all trajectories eventually remain. By using a feedback $u = \rho x_1$ with large ρ, the diameter of the ellipsoid can be scaled by $1/\rho$ without losing those properties. □

Čelikovský and Vaněček [51] points out that the famous Lorenz [193] dynamical system can be represented as a bilinear control system

$$\dot{x} = Ax + uBx,\ A = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho & -1 & 0 \\ 0 & 0 & -\beta \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.24}$$

with positive parameters σ, ρ, β and linear feedback $u = x_1$. For parameter values near the nominal $\sigma = 10, b = 8/3, \rho = 28$, the Lorenz system has a bounded limit-set of unstable periodic trajectories with indefinitely long periods (a strange attractor); for pictures see [51], which shows that such chaotic behavior is common if, as in (3.24), $\dot{x} = Bx$ is a rotation and A has real eigenvalues, of which two are negative and one positive. In (3.24) that occurs if $\rho > 1$.

For the Lorenz system (3.24), $\det(A + sB) = b\sigma(\rho - 1) - \sigma s^2$, whose roots $\pm\sqrt{b(\rho - 1)}$ are real for $\rho \geq 1$. The two corresponding equilibrium points are

$$\mathrm{col}\left(\sqrt{b(\rho-1)}, \sqrt{b(\rho-1)}, \rho - 1\right),\ \mathrm{col}\left(-\sqrt{b(\rho-1)}, -\sqrt{b(\rho-1)}, \rho - 1\right).$$

Next, a computation with the script of Table 2.4 reveals that $\ell = 9$, so $\{A, B\}_{\mathcal{L}} = \mathfrak{gl}(3, \mathbb{R})$ and (supposing $\rho\sigma > 0$) has radical ideal $\mathcal{I} = \langle x_3, x_2, x_1 \rangle$.

Example 3.14. The rotated semistable systems are not sufficiently general to deal with several other well-known quadratic systems that show bounded chaos. There are biaffine systems with feedback $u = x_1$ that have chaotic behavior including the Rössler attractor

$$\dot{x} = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 3/20 & 0 \\ 0 & 0 & -10 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1/5 \end{bmatrix} + x_1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} x$$

and self-excited dynamo models of the Earth's magnetic fields [126] and other homopolar electric generators

$$\dot{x} = \begin{bmatrix} -1 & 0 & -\beta \\ -\alpha & -\kappa & 0 \\ 1 & 0 & -\lambda \end{bmatrix} x + \begin{bmatrix} 0 \\ \alpha \\ 0 \end{bmatrix} + x_1 \begin{bmatrix} 0 & 1 & 0 \\ \alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} x. \qquad \triangle$$

That any polynomial dynamical system $\dot{x} = Ax + p(x)$ with $p(\lambda x) = \lambda^2 q(x)$ can be obtained by using linear feedback in some bilinear system (3.1) was mentioned in Bruni et al. [43] and an explicit construction was shown by Frayman [96, 97]. Others including Markus [200] and Kinyon and Sagle [160] have studied homogeneous quadratic dynamical systems $\dot{x} = q(x)$. Let $\beta(x, y) := q(x + y) - q(x) - q(y)$. A nonassociative commutative algebra $\mathfrak{Q}$ is defined using the multiplication $x * y = \beta(x, y)$. Homogeneous quadratic dynamical systems may have multiple equilibria, finite escape time (see Section B.3.2) limit cycles, and invariant cones; many qualitative properties are the dynamical consequences of various algebraic properties (semisimplicity, nilpotency, etc.) of $\mathfrak{L}$.

3.7 Lie Semigroups

The controllability of (3.1) on $\mathbb{R}^n_*$ has become a topic in the considerable mathematical literature on Lie semigroups.[23] The term Lie semigroup will be used to mean a subsemigroup (with identity) of a matrix Lie group. The semigroup **S** transition matrices of (3.2) is almost the only kind of Lie semigroup that will be encountered here, but there are others. For introductions to Lie theory relevant to control theory, see the foundational books Hilgert et al. [127] and Hilgert and Neeb [128]. For a useful exposition, see Lawson [173]; there and in many mathematical works the term "bilinear system" refers to (3.2) evolving on a *closed Lie subgroup* $\mathbf{G} \subset GL^+(n, \mathbb{R})$.[24]

We begin with the fact that for system (3.2) *acting as a control system on* **G**, a matrix Lie group whose Lie algebra is $\mathfrak{g} = \{A, B_1, \ldots, B_m\}_{\mathcal{L}}$, the attainable set is $\mathbf{S}_\Omega \subset \mathbf{G} \subset GL^+(n, \mathbb{R})$. If **G** is closed, it is an embedded submanifold of $GL^+(n, \mathbb{R})$; in that case, we can regard (3.2) as a right-invariant control system on **G**.

For (3.1) to be controllable on $\mathbb{R}^n_*$, it is first necessary (but far from sufficient) to show that $\mathfrak{g}$ is a transitive Lie algebra from the list in Appendix D; that means the corresponding Lie subgroup of $GL^+(n, \mathbb{R})$ is transitive on $\mathbb{R}^n_*$. If $\mathbf{S} = \mathbf{G}$ then (3.1) is controllable. In older works **G** was either $SL(n, \mathbb{R})$ or $GL^+(n, \mathbb{R})$, whose Lie algebras are easily identified by the dimension of their bases ($\ell = n - 1$ or $\ell = n$) obtained by the `LieTree` algorithm in Chapter 2. Recently do Rocio et al. [75] carried out this program for subsemigroups of each of the transitive Lie groups listed in Boothby and Wilson [32].

The important work of Jurdjevic and Kupka [148, Parts II–III] (also [147]) on matrix control systems (3.2) defines the accessibility set to be $\overline{\mathbf{S}}$, the topological closure of **S** in **G**. The issue they considered is: when is $\overline{\mathbf{S}} = \mathbf{G}$?

Sufficient conditions for this equality were given in [148] for connected simple or semisimple Lie groups **G**. Two control systems on **G** are called *equivalent* in [148] if their accessibility sets are equal. LS called the Lie saturate of (3.2) in Jurdjevic and Kupka [148].

---

[23] For right-invariant control systems (3.2) on closed Lie groups, see the basic papers of Brockett [36] and Jurdjevic and Sussmann [151], and many subsequent studies such as [37, 127, 128, 147, 148, 180, 254]. Sachkov [230, 232] are introductions to Lie semigroups for control theorists.

[24] Control systems like $\dot{X} = [A, X] + uBX$ on closed Lie groups have a resemblance to linear control systems and have been given that name in Ayala and Tirao [12] and Cardetti and Mittenhuber [48], which discuss their controllability properties.

$$\mathrm{LS} := \{A, B_1, \ldots, B_m\}_{\mathcal{L}} \bigcap \{X \in \mathcal{L}(\mathbf{G}) | \exp(tX) \in \overline{\mathbf{S}} \text{ for all } t > 0\} \tag{3.25}$$

is called the Lie saturate of (3.2). Example 2.5 illustrates the possibility that $\mathbf{S} \neq \overline{\mathbf{S}}$.

Proposition 3.14 (Jurdjevic and Kupka [148, Prop. 6]). *If* LS $= \mathcal{L}(\mathbf{G})$, *then* (3.2) *is controllable.* △

Theorem 3.11 (Jurdjevic and Kupka [148]). *Assume that* $\mathrm{tr}(A) = 0 = \mathrm{tr}(B)$ *and* B *is strongly regular. Choose coordinates so that* $B = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$. *If* A *satisfies*

$$a_{i,j} \neq 0 \textit{ for all } i, j \textit{ such that } |i - j| = 1 \textit{ and} \tag{3.26}$$

$$a_{1,n}a_{n,1} < 0 \tag{3.27}$$

then with $\Omega = \mathbb{R}$, $\dot{x} = (A + uB)x$ *is controllable on* $\mathbb{R}^n_*$.

Definition 3.9. A matrix $P \in \mathbb{R}^{n\times n}$ is called a permutation matrix if it can be obtained by permuting the rows of I_n. A square matrix A is called permutation-reducible if there exists a permutation matrix P such that

$$P^{-1}AP = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} \text{ where } A_3 \in \mathbb{R}^{k\times k},\ 0 < k < n.$$

If there exists no such P then[25] P is called permutation-irreducible. △

Gauthier and Bornard [102, Th.3] shows that in Theorem 3.11 the hypothesis (3.26) on A can be replaced by the assumption that A is permutation-irreducible.

Hypothesis (3.26) can also be replaced by the condition that $\{A, B\}_{\mathcal{L}} = \mathfrak{sl}(n, \mathbb{R})$, which was used in a new proof of Theorem 3.11 in San Martin [233]; that condition is obviously necessary.

For two-dimensional bilinear systems, the semigroup approach has produced an interesting account of controllability in Braga Barros et al. [23]. This paper first gives a proof that there are only three connected Lie groups transitive on $\mathbb{R}^2_*$; they are $\mathrm{GL}^+(2, \mathbb{R})$, $\mathrm{SL}(2, \mathbb{R})$, $\mathbb{C}_* \simeq \mathrm{SO}(2) \times \mathbb{R}^1_+$, as in Appendix D. These are discussed separately, but the concluding section gives necessary and sufficient conditions[26] for the controllability of (3.5) on R^2_*, paraphrased here for reference.

- Controllability $\Leftrightarrow$ $\det([A, B]) < 0$ in each of these cases:

  1. $\det(B) \leq 0$ and $B \neq 0$.
  2. $\det(B) > 0$ and $\mathrm{tr}(B)^2 - 4\det(B) > 0$.
  3. $\mathrm{tr}(A) = 0 = \mathrm{tr}(B)$.

[25] There are graph-theoretical tests for permutation-reducibility, see [102].

[26] Some of these conditions, it is mentioned in [23], had been obtained in other ways by Lepe [181] and by Joó and Tuan [145].

- If $B \simeq \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, then controllability $\Leftrightarrow$ $\det([A, B]) \neq 0$.
- If $B = kI$, then controllability $\Leftrightarrow$ $\mathrm{tr}(A)^2 - 4\det(A) < 0$.
- If $\det(B) > 0$ and $\mathrm{tr}(B) \neq 0$, then controllability $\Leftrightarrow$ A and B are linearly independent.
- If B is neutral, then controllability $\Leftrightarrow$ $\det([A, B]) + \det(B)\,\mathrm{tr}(A)^2 < 0$.
- If B is neutral and $\mathrm{tr}(A) = 0$, then controllability $\Leftrightarrow$ $[A, B] \neq 0$.

Ayala and San Martin [11] also contributed to this problem in the case that A and B are in $\mathfrak{sl}(2, \mathbb{R})$.

3.7.1 *Sampled Accessibility and Controllability**

Section 1.8(1) described the sampled-data control system

$$x^+ = \mathrm{e}^{\tau(A+uB)}x, \;\; u(\cdot) \in \Omega; \tag{3.28}$$
$$\mathbf{S} := \{e^{\tau(A+u(k)B)} \dots e^{\tau(A+u(1)B)} \,|\, k \in \mathbb{N}\}$$

is its transition semigroup. Since it is linear in x, it provides easily computed solutions of $\dot{x} = Ax + uBx$ every τ seconds.

System (3.28) is said to have the sampled accessibility property if for all ξ there exists $T \in \mathbb{N}$ such that

$$\{e^{\tau(A+u(T)B)} \dots e^{\tau(A+u(1)B)}\xi \,|\, u(\cdot) \in \Omega\}$$

has open interior. For the corresponding statement about sampled observability, see Proposition 5.3.

Proposition 3.15 (Sontag [248]). *If* $\dot{x} = (A+uB)x$ *has the accessibility property on* $\mathbb{R}^n_*$ *and every quadruple* $\lambda, \mu, \lambda', \mu'$ *of eigenvalues of* A *satisfies* $\tau(\lambda + \lambda' - \mu - \mu') \neq 2k\pi,\; k \neq 0$ *then* (3.28) *will have the sampled accessibility property.* △

If $\{A, B\}_{\mathcal{L}} = \mathfrak{sl}(n, \mathbb{R})$ and the pair (A, B) satisfies an additional regularity hypothesis (stronger than that of Theorem 3.11), then [233, Th. 3.4, 4.1] $\mathbf{S} = \mathrm{SL}(n, \mathbb{R})$, which is transitive on $\mathbb{R}^n_*$; the sampled system is controllable.

3.7.2 *Lie Wedges**

Remark 2.3 pointed out that the set $\{X \in \mathfrak{gl}(n, \mathbb{R}) \,|\, \exp(\mathbb{R}X) \in \mathbf{G}\}$, the tangent space of a matrix Lie group $\mathbf{G}$ at the identity, is its Lie algebra $\mathfrak{g}$. When is there an analogous fact for a transition matrix semigroup $\mathbf{S}_\Omega$? There is a sizable literature on this question, for more general semigroups. The chapter by J. Lawson [173] in [89] is the main source for this section.

A subset $\mathbf{W}$ of a matrix Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ is called a closed wedge if $\mathbf{W}$ is closed under addition and by multiplication by scalars in $\mathbb{R}_+$, and is topologically closed in $\mathfrak{g}$. (A wedge is just a closed convex cone, but in Lie semigroup theory the new term is useful.) The set $\mathbf{W} \cup -\mathbf{W}$ is largest subspace contained in $\mathbf{W}$ and is called the edge of $\mathbf{W}$. If the edge is a subalgebra of $\mathfrak{g}$, it is called a Lie wedge.

Definition 3.10. The subtangent set of $\mathbf{S}$ is

$$\mathbf{L}(\mathbf{S}) := \{X \in \mathfrak{g} | \ \exp \mathbb{R}_+ X \in \mathbf{S}\}.$$

The inclusion $\mathbf{L}(\mathbf{S}) \subset \mathfrak{g}$ is proper unless $A \in \operatorname{span}\{B_1, \dots, B_m\}$. △

Theorem 3.12 (Lawson [173, Th. 4.4]). *If* $\mathbf{S} \subseteq \mathbf{G}$ *is a closed subsemigroup containing* I*, then* $\mathbf{L}(\mathbf{S})$ *is a Lie wedge.*

However, given a wedge $\mathbf{W}$ there is, according to [173], no easy route to a corresponding Lie semigroup. For introductions to Lie wedges and Lie semigroups, Hilgert et al. [127] and Hilgert and Neeb [128] are notable. Their application to control on solvable Lie groups was studied by Lawson and Mittenhuber [174]. For the relationship between Lie wedge theory and the Lie saturates of Jurdjevic and Kupka, see Lawson [173].

3.8 Biaffine Systems

Many of the applications mentioned in Chapter 6 lead to the biaffine (inhomogeneous bilinear) control systems introduced in Section 1.4.[27]

This section is primarily about some aspects of the system

$$\dot{x} = Ax + a + \sum_{i=1}^{m} u_i(t)(B_i x + b_i), \quad x, a, b_i \in \mathbb{R}^n \tag{3.29}$$

and its discrete-time analogs that can be studied by embedding them as homogeneous bilinear systems in $\mathbb{R}^{n+1}$ as described below in Sections 3.8.1 and 4.4.6, respectively. Section 3.8.3 contains some remarks on stabilization.

The controllability theory of (3.29) began with the 1968 work of Rink and Mohler [224], [206, Ch. 2], where the term "bilinear systems" was first introduced. Their method was strengthened in Khapalov and Mohler [158];

---

[27] Mohler's books [205–207] contain much about biaffine systems — controllability, optimal control, and applications, using control engineering methods. The general methods of stabilization given in Isidori [140, Ch. 7] can be used for biaffine systems for which the origin is not a fixed point, and are beyond the scope of this book.

the idea is to find controls that provide $2m+1$ equilibria of (3.29) at which the local uncontrolled linear dynamical systems have hyperbolic trajectories that can enter and leave neighborhoods of these equilibria (approximate control) and to find one equilibrium at which the Kalman rank condition holds.

Biaffine control systems can be studied by a Lie group method exemplified in the 1982 paper of Bonnard et al. [28], whose main result is given below as Theorem 3.13, and in subsequent work of Jurdjevic and Sallet quoted as Theorem 3.14. These are prefaced by some definitions and examples.

3.8.1 Semidirect Products and Affine Groups

Note that $\mathbb{R}^n$, with addition as its group operation, is an Abelian Lie group. If $\mathbf{G} \subset \mathrm{GL}(n, \mathbb{R})$, then the semidirect product of $\mathbf{G}$ and $\mathbb{R}^n$ is a Lie group $\widehat{\mathbf{G}} := \mathbb{R}^n \ltimes \mathbf{G}$ whose underlying set is $\{(v, X) | v \in \mathbb{R}^n, X \in \mathbf{G}\}$ and whose group operation is $(v, X) \cdot (w, Y) := (v + Xw, XY)$ (see Definition B.16 for semi-direct products). Its standard representation on $\mathbb{R}^{n+1}$ is

$$\widehat{\mathbf{G}} = \left\{ \widehat{A} := \begin{bmatrix} A & a \\ \mathbf{0}_{1,n} & 1 \end{bmatrix} \Big| A \in \mathbf{G}, a \in \mathbb{R}^n \right\} \subset \mathrm{GL}(n+1, \mathbb{R}).$$

The Lie group of all affine transformations $x \to Ax + a$ is $\mathbb{R}^n \ltimes \mathrm{GL}(n, \mathbb{R})$, whose standard representation is called $\mathbf{Aff}(n, \mathbb{R})$. The Lie algebra of $\mathbf{Aff}(n, \mathbb{R})$ is the representation of $\mathbb{R}^n \dotplus \mathfrak{gl}(n, \mathbb{R})$ (see Definition B.16) and is explicitly

$$\mathfrak{aff}(n, \mathbb{R}) := \left\{ \widehat{X} := \begin{bmatrix} X & x \\ \mathbf{0}_{1,n} & 0 \end{bmatrix} \Big| X \in \mathfrak{gl}(n, \mathbb{R}),\ x \in \mathbb{R}^n \right\} \subset \mathfrak{gl}(n, \mathbb{R}^{n+1}); \tag{3.30}$$

$$[\widehat{X}, \widehat{Y}] = \left[\begin{bmatrix} X & x \\ \mathbf{0}_{1,n} & 0 \end{bmatrix}, \begin{bmatrix} Y & y \\ \mathbf{0}_{1,n} & 0 \end{bmatrix} \right] = \begin{bmatrix} [X, Y] & Xy - Yx \\ \mathbf{0}_{1,n} & 0 \end{bmatrix}.$$

With the help of this representation, one can easily use the two canonical projections

$$\pi_1 : \widehat{\mathbf{G}} \to \mathbb{R}^n,\ \pi_2 : \widehat{\mathbf{G}} \to \mathbf{G},$$

$$\pi_1 \begin{bmatrix} A & a \\ \mathbf{0}_{1,n} & 1 \end{bmatrix} = a,\ \pi_2 \begin{bmatrix} A & a \\ \mathbf{0}_{1,n} & 1 \end{bmatrix} = A.$$

For example, $\mathbb{R}^n \ltimes \mathrm{SO}(n) \subset \mathbf{Aff}(n, \mathbb{R})$ is the group of translations and proper rotations of Euclidean n-space. For more examples of this kind, see Jurdjevic [147, Ch. 6].

Using the affine group permits us to represent (3.29) as a bilinear system whose state vector is $z := \mathrm{col}(x_1, \ldots, x_n, x_{n+1})$ and for which the hyperplane

$L := \{z|\, x_{n+1} = 1\}$ is invariant. This system is

$$\dot{z} = \widehat{A}z + \sum_{i=1}^{m} u_i \widehat{B}_i z, \quad \text{for } \widehat{A} = \begin{bmatrix} A & a \\ \mathbf{0}_{1,n} & 0 \end{bmatrix}, \ \widehat{B}_i = \begin{bmatrix} B_i & b_i \\ \mathbf{0}_{1,n} & 0 \end{bmatrix}. \tag{3.31}$$

The Lie algebra of the biaffine system (3.31) can be obtained from its Lie tree, which for $m = 1$ is $\mathcal{T}(\widehat{A}, \widehat{B})$; for example

$$[\widehat{A}, \widehat{B}] = \begin{bmatrix} [A, B] & Ab - Ba \\ \mathbf{0}_{1,n} & 0 \end{bmatrix},$$

$$[\widehat{A}, [\widehat{A}, \widehat{B}]] = \begin{bmatrix} [A, [A, B]] & A^2 b - ABa - [A, B]a \\ \mathbf{0}_{1,n} & 0 \end{bmatrix},$$

$$[\widehat{B}, [\widehat{B}, \widehat{A}]] = \begin{bmatrix} [B, [B, A]] & B^2 a - BAb - [B, A]b \\ \mathbf{0}_{1,n} & 0 \end{bmatrix}, \ \ldots .$$

A single-input linear control system[28] $\dot{x} = Ax + ub$ corresponds to a special case of (3.29). For its representation in $\mathbf{Aff}(n, \mathbb{R})$, let

$$z := \operatorname{col}(x_1, \ldots, x_n, x_{n+1}); \ \dot{z} = \widehat{A}z + u(t)\widehat{B}z, \ z(0) = \begin{bmatrix} \xi \\ 1 \end{bmatrix}, \text{ where}$$

$$\widehat{A} = \begin{bmatrix} A & \mathbf{0}_{n,1} \\ \mathbf{0}_{1,n} & 0 \end{bmatrix}, \ \widehat{B}_i = \begin{bmatrix} \mathbf{0}_{n,n} & b \\ \mathbf{0}_{1,n} & 0 \end{bmatrix}; \ \operatorname{ad}^k_{\widehat{A}}(\widehat{B}) = \begin{bmatrix} \mathbf{0}_{n,n} & A^k b \\ \mathbf{0}_{1,n} & 0 \end{bmatrix}.$$

Corresponding to the projection π_1 of groups, there is a projection dp_1 of the Lie algebra $\{\widehat{A}, \widehat{B}\}_{\mathcal{L}} \subset \mathfrak{aff}(n, \mathbb{R})$:

$$d\pi_1 \left\{ \widehat{A}z, \widehat{B}z, \operatorname{ad}^1_{\widehat{A}}(\widehat{B})z, \ \ldots, \operatorname{ad}^n_{\widehat{A}}(\widehat{B})z \right\} = \{Ax, b, Ab, \ldots, A^{n-1}b\}, \tag{3.32}$$

$$\operatorname{rank}\left[Ax \ b \ Ab \ \ldots \ A^{n-1}b \right] = n \text{ for all } x \in \mathbb{R}^n$$

$$\Leftrightarrow \quad \operatorname{rank}\left[b \ Ab \ \ldots \ A^{n-1}b \right] = n,$$

which is the ad-condition for $\dot{x} = Ax + ub$.

3.8.2 *Controllability of Biaffine Systems*

Bonnard et al. [28] dealt with semidirect products $\widehat{\mathbf{G}} := V \rtimes \mathbf{K}$ where $\mathbf{K}$ is a compact Lie group with Lie algebra $\mathfrak{k}$ and we can take $V = \mathbb{R}^n$; the Lie algebra of $\widehat{\mathbf{G}}$ is $\mathbb{R}^n \dotplus \mathfrak{k}$ and its matrix representation is a subalgebra of $\mathfrak{aff}(n, \mathbb{R})$. The

[28] This way of looking at linear control systems as bilinear systems was pointed out in Brockett [36]. The same approach to multiple-input linear control systems is an easy generalization, working on $\mathbb{R}^{n+m}$.

system (3.31) acts on (the matrix representation of) $\widehat{\mathbf{G}}$. Let $\widehat{\mathfrak{g}} = \{\widehat{A}, \widehat{B}_1, \dots, \widehat{B}_m\}_{\mathcal{L}}$ and let $\mathbf{S}$ be the semigroup of transition matrices of (3.31). The first problem attacked in [28] is to find conditions for the transitivity of $\mathbf{S}$ on $\mathbf{G}$; we will discuss the matrix version of this problem. One says that x is a fixed point of the action of $\mathbf{K}$ if $\mathbf{K}x = x$ for some $x \in \mathbb{R}^n$.

Theorem 3.13 (BJKS [28, Th. 1]). *Suppose that* 0 *is the only fixed point of the compact Lie group* $\mathbf{K}$ *on* $\mathbb{R}^n$. *Then for* (3.31), $\mathbf{S}$ *is transitive on* $\widehat{\mathbf{G}}$ *if and only if* $\widehat{\mathfrak{g}} = \mathbb{R}^n \dotplus \mathfrak{k}$.

The LARC on $\widehat{\mathbf{G}}$ projects to the LARC on $\mathbf{K}$, so transitivity of $\mathbf{S}$ on $\mathbf{K}$ follows from Proposition 3.9.

To apply the above theorem to (3.29), following [28], suppose that the bilinear system on $\mathbf{K}$ induced by π_2 is transitive on $\mathbf{K}$, the LARC holds on $\mathbb{R}^n_*$ and that $\mathbf{K}x = 0$ implies $x = 0$; then (3.29) is controllable on $\mathbb{R}^n$.

Jurdjevic and Sallet [150, Th. 2, 4] in a similar way obtain sufficient conditions for the controllability of (3.29) on $\mathbb{R}^n$.

Theorem 3.14 (Jurdjevic and Sallet). *Given*

(I) the controllability on $\mathbb{R}^n_*$ *of the homogeneous system*

$$\dot{x} = Ax + \sum_{i=1}^{m} u_i(t)B_i x, \ \Omega = \mathbb{R}^m; \textit{ and} \tag{3.33}$$

(II) for each $x \in \mathbb{R}^n$ *there exists some constant control* $u \in \mathbb{R}^m$ *for which* $\dot{X} \neq 0$ *for* (3.29) *(the* no-fixed-point *condition);*

then the inhomogeneous system (3.29) *is controllable on* $\mathbb{R}^n$*; and for sufficiently small perturbations of its coefficients, the perturbed system remains controllable.*

Example 3.15. A simple application of Theorem 3.13 is the biaffine system on $\mathbb{R}^2$ described by $\dot{x} = uJx + a$, $\Omega = \mathbb{R}$, $a \neq 0$. The semidirect product here is $\mathbb{R}^2 \ltimes T_1$. Since T^1 is compact; $0 \in \mathbb{R}^2$ is its only fixed point; and the Lie algebra

$$\left\{\begin{bmatrix} 0 & u & a_1 \\ -u & 0 & a_2 \\ 0 & 0 & 0 \end{bmatrix}\right\}_{\mathcal{L}} = \operatorname{span}\left\{\begin{bmatrix} 0 & u_1 & u_2a_1 + u_3a_2 \\ -u_1 & 0 & u_3a_2 - u_2a_1 \\ 0 & 0 & 0 \end{bmatrix}\right\}$$

is the Lie algebra of $\mathbb{R}^2 \ltimes T_1$, controllability follows from the Theorem. Note: the trajectories are given by $x(t) = \exp(Jv(t))\xi + \Lambda(J; v(t))a$, where $v(t) = \int_0^t u(s)\,ds$ and Λ is defined in Section 1.1.3; using (1.14), $x(T) = \xi$ for any T and u such that $v(T) = 2\pi$. △

Example 3.16. Suppose that the bilinear system $\dot{x} = Ax + uBx$ satisfies $A^\top + A = 0$ and the ad-condition on $\mathbb{R}^n_*$; it is then small-controllable. Now consider the biaffine system with the same matrices, $\dot{x} = Ax + a + u(Bx + b)$. Impose the no-fixed-point condition: the equations $Ax + a = 0, Bx + b = 0$ are inconsistent. Then from Theorem 3.14, the biaffine system is controllable on $\mathbb{R}^n$. △

3.8.3 Stabilization for Biaffine Systems

In studying the stabilization of $\dot{x} = f(x) + ug(x)$ at the origin by a state feedback $u = \phi(x)$, the standard assumption is that $f(0) = 0$, $g(0) \neq 0$, to which more general cases can be reduced. Accordingly, let us use the system $\dot{x} = Ax + (Bx + b)u$ as a biaffine example. The chief difficulty is with the set $E(u) := \{x| \, Ax + u(Bx + b) = 0\}$; there may be equilibria whose location depends on u. We will avoid the difficulty by making the assumption that for all u we have $E(u) = \emptyset$. The set $\{A, B, b\}$ that satisfies this no-fixed-point assumption is generic (as defined in Remark 2.5) in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \times \mathbb{R}^n$. A second generic assumption is that (A, b) is a controllable pair; then it is well known (see Exercises 1.11 and 1.14) that one can readily choose $c^\top$ such that $A + c^\top b$ is a Hurwitz matrix. Then the linear feedback $u = c^\top x$ results in the asymptotic stability of the origin for $\dot{x} = (A + bc^\top)x + (c^\top x)Bx$ because the quadratic term can be neglected in a neighborhood of the origin. Finding the basin of stability is not a trivial task. For $n = 2$, if the basin is bounded, under the no-fixed-point assumption the Poincaré-Bendixson Theorem shows that the basin boundary is a limit cycle, not a separatrix. In higher dimensions, there may be strange attractors.

There is engineering literature on piecewise constant sliding-mode control to stabilize biaffine systems, such as Longchamp [192] which uses the sliding-mode methods of Krasovskii [163].

Exercise 3.6. Show (using $V(x) := x^\top x$) that the origin is globally asymptotically stable for the system

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} x + u \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x + u \begin{bmatrix} 0 \\ -1 \end{bmatrix} \text{ if } u = x_2. \qquad \triangle$$

3.8.4 Quasicommutative Systems

As in Section 3.8.1, let the biaffine system (with $a = 0$)

$$\dot{x} = Ax + (Bx + b)u, \; x \in \mathbb{R}^n, \text{ be represented on } \mathbb{R}^{n+1} \text{ by} \tag{3.34}$$

$$\dot{z} = \widehat{A}z + u(t)\widehat{B}z, \; z = \begin{bmatrix} x \\ z_{n+1} \end{bmatrix}, \; z_{n+1}(0) = 1. \tag{3.35}$$

If the matrices $\widehat{B}, \mathrm{ad}^1_{\widehat{A}}(\widehat{B}), \dots, \mathrm{ad}^{n-1}_{\widehat{A}}(\widehat{B})$ all *commute* with each other, the pair $(\widehat{A}, \widehat{B})$ and the system (3.35) are called quasicommutative in Čelikovský [50], where it is shown that this property characterizes linearizable control systems (3.34) among biaffine control systems. If (3.35) is quasicommutative and there exists a state ξ for which $(A, B\xi + b)$ is a controllable pair, then by [50, Th. 5] there exists a smooth diffeomorphism ϕ on a neighborhood of ξ

that takes (3.34) to a controllable system $\dot{y} = Fy + ug$ where $F \in \mathbb{R}^{n\times n}$, $g \in \mathbb{R}^n$ are constants. The projection $\mathbb{R}^{n+1} \to \mathbb{R}^n$ that kills x_{n+1} is represented by the matrix $P_n := \begin{bmatrix} I_n & \mathbf{0}_{1,n} \end{bmatrix}$. Define

$$\psi(y) := P_n \exp\left(\sum_{k=0}^{n-1} y_{k+1} \operatorname{ad}_A^k(B) \right) z(0).$$

Example 3.17 (Čelikovský [50]). For (3.34) with $n = 3$ let $\xi := 0$,

$$A := \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \; b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \; \operatorname{rank}(b, Ab, A^2 b) = 3.$$

$$\psi(y) = \sum_{k=0}^{2} \begin{bmatrix} 3^k \\ 2^k \\ 1 \end{bmatrix} y_{k+1} + \frac{1}{2}(y_1 + y_2 + y_3)^2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

is the required diffeomorphism.

Exercise: show that ψ^{-1} is quadratic and that y satisfies $\dot{y} = Fy + ug$, where (F, g) is a controllable pair. △

3.9 Exercises

Exercise 3.7. For Section 3.6.3, with arbitrary $Q \gg 0$ find the minima and maxima over $\mathbb{R}^n_*$ of

$$\frac{c^\top x}{\sqrt{x^\top Q x}} \text{ and } \frac{x^\top R x}{x^\top Q x}.$$ △

Exercise 3.8. If $n = 2$, show that A is neutral if and only if $\operatorname{tr}(A) = 0$ and $\det(A) > 0$. △

Exercise 3.9. For the Examples 1.7 and 1.8 in Chapter 1, find the rank of $M(x)$, of $M_a(x)$, and of $M_a^w(x)$. △

Exercise 3.10. The following Mathematica script lets you verify Lemma 3.2 for any $1 \le d < n$, $1 \le k, m \le n$. The case $d = n$, where `Minors` fails, is just $\det(PQ) = \det(P)\det(Q)$.

```
 n = 4; P = Table[p[i, j], i, n, j, n];
Q = Table[q[i, j], i, n, j, n]; R = P . Q;
d = 2; c = n!/(d!*(n - d)!); k = 4; m = 4;
Rminor = Expand[Minors[R, d][[k,m]]];
sumPQminors =
    Expand[Sum[Minors[P,d][[k,i]]*Minors[Q,d][[i,m]],i,c]];
Rminor - sumPQminors
```
△

Exercise 3.11. The Mathematica script below is set up for Example 3.12. To choose a denominator den for u, enter 1 or 2 when the program asks for denq. The init initial states lie equispaced on a circle of radius amp. With a few changes (use a new copy) this script can be used for Example 3.13 and its choices of ϕ.

```
 A={{0,-1},{1,0}};B={{1,0},{0,-1}}; X={x,y};
u=(-x/5-y)/den; denq=Input["denq"];
den=Switch[denq, 1,1, 2,Sqrt[X.X]];
F=(A+u*B).X/.{x->x[t],y->y[t]};
Print[A//MatrixForm,B//MatrixForm]; Print["F= ",F]
time=200; inits=5; amp=0.44;
For[i=1,i<=inits,i++,tau=2*Pi/inits;
S[tau_]:=amp*{Cos[i*tau],Sin[i*tau]};
eqs[i]={x'[t]==F[[1]],y'[t]==F[[2]],
x[0]==S[tau][[1]],y[0]==S[tau][[2]]};
sol[i]=NDSolve[eqs[i],x,y,t,0,time,MaxSteps->20000]];
res=Table[x[t],y[t]/.sol[i],{i,inits}];
Off[ParametricPlot::"ppcom"];
ParametricPlot[Evaluate[res],{t,0,time},PlotRange->All,
PlotPoints->500,AspectRatio->Automatic,Axes->True];
Plot[Evaluate[(Abs[x[t]]+Abs[y[t]])/.sol[inits]],t,0,time];
(* The invariants under similarity of the pair (A,B): *)
i0=Det[B];
i1=(Tr[A]*Tr[B]-Tr[A.B])^2-4*Det[A]*Det[B];
i2=Tr[A]*Tr[B]-Tr[A.B]*Tr[B.B];
i3=2Tr[A]*Tr[B]*Tr[A.B]-Tr[A.A]*Tr[B]^2-Tr[A]^2*Tr[B.B];
CK[A1_,B1_]:=4*Tr[A1.B1]-2*Tr[A1]*Tr[B1];
i4=CK[A,B]^2-CK[A,A]*CK[B,B];
Simplify[{i0,i1,i2,i3,i4}]
```

△

Chapter 4
Discrete-Time Bilinear Systems

This chapter considers the discrete-time versions of some of the questions raised in Chapters 2 and 3, but the tools used are different and unsophisticated. Similar questions lead to serious problems in algebraic geometry that were considered for general polynomial discrete-time control systems in Sontag [244] but are beyond the scope of this book. Discrete-time bilinear systems were mentioned in Section 1.8 as a method of approximating the trajectories of continuous-time bilinear systems. Their applications are surveyed in Chapter 6; observability and realization problems for them will be discussed in Chapter 5 along with the continuous-time case.

Contents of this Chapter

Some facts about discrete-time linear dynamical systems and control systems are sketched in Section 4.1, including some basic facts about stability. The stabilization problem introduced in Chapter 3 has had less attention for discrete-time systems; the constant control version has a brief treatment in Section 4.3.

Discrete-time control systems are first encountered in Section 4.2; Section 4.4 is about their controllability. To show noncontrollability, one can seek invariant polynomials or invariant sets; some simple examples are shown in Section 4.4.1. If the rank of B is one, it can be factored, $B = bc^\top$, yielding a class of multiplicative control systems that can be studied with the help of a related linear control system as in Section 4.4.2.

Small-controllability was introduced in Section 3.5; the discrete-time version is given in Section 4.4.3. The local rank condition discussed here is similar to the *ad*-condition, so in Section 4.4.5 non-constant stabilizing feedbacks are derived. Section 4.5 is an example comparing a noncontrollable continuous-time control system with its controllable Euler discretization.

D.L. Elliott, *Bilinear Control Systems*, Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_4, © Springer Science+Business Media B.V. 2009

4.1 Dynamical Systems: Discrete-Time

A real discrete-time linear dynamical system $\overset{+}{x} = Ax$ can be described as the action on $\mathbb{R}^n$ of the semigroup $\mathbf{S}_A := \{A^t \mid t \in \mathbb{Z}_+\}$.

If the Jordan canonical form of a real matrix A meets the hypotheses of Theorem A.3, a real logarithm $F = \log(A)$ exists.

Proposition 4.1. *If $F = \log(A)$ exists, $\overset{+}{x} = Ax$ becomes $x(t) = \exp(tF)\xi$ and the trajectories of $\overset{+}{x} = Ax$ are discrete subsets of the trajectories of $\dot{x} = Ax$:*

$$\mathbf{S}_A \subset \{\exp(\mathbb{R}_+ F)\}. \qquad \triangle$$

If for some $k \in \mathbb{Z}_+$ $A^k = I$ then $\mathbf{S}_A$ is a finite group. If A is singular there exists $C \neq 0$ such that $CA = 0$, and for any initial state $x(0)$ the trajectory satisfies $Cx(t) = 0$. If A is nilpotent (i.e., A is similar to a strictly upper-triangular matrix), there is some $k < n$ such that trajectories arrive at the origin in at most k steps.

For each eigenvalue $\mu \in \operatorname{spec}(A)$, there exists ξ for which $\|x(t)\| = |\mu|^t |\xi|$. If $\operatorname{spec}(A)$ lies inside the unit disc $D \subset \mathbb{C}$, all trajectories of $\overset{+}{x} = Ax$ approach the origin and the system is strictly stable. If even one eigenvalue μ lies outside D, there exist trajectories that go to infinity and the system is said to be unstable. Eigenvalues on the unit circle (neutral stability) correspond in general to trajectories that lie on an ellipsoid (sphere for the canonical form) with the critical or resonant case of growth polynomial in k occurring when any of the Jordan blocks is one of the types

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \cdots .$$

To discuss stability criteria for the dynamical system $\overset{+}{x} = Ax$ in $\mathbb{R}^n$, we need a discrete-time Lyapunov method. Take as a gauge function $V(x) := x^\top Q x$, where we seek $Q \gg 0$ such that

$$W(x) := V(\overset{+}{x}) - V(x) = x^\top (A^\top Q A - Q) x \ll 0$$

to show that trajectories all approach the origin. The discrete-time stability criterion is that there exists a positive definite solution Q of

$$A^\top Q A - Q = -I. \tag{4.1}$$

To facilitate calculations, (4.1) can be rewritten in vector form. As in Exercises 1.6–1.8, we use the Kronecker product $\otimes$ of Section A.3; $\flat : \mathbb{R}^{n \times n} \to \mathbb{R}^\nu$ is the flattening operator that takes a square matrix to a column of its ordered columns. Using (A.9),

$$A^\top X A - X = -P \;\Leftrightarrow (A^\top \otimes A^\top - I \otimes I)X^\flat = -P^\flat \tag{4.2}$$

in $\mathbb{R}^\nu$. From Section A.3.1 and using $I_n \otimes I_n = I_\nu$, the ν eigenvalues of $A^\top \otimes A^\top - I \otimes I)$ are $\{\alpha_i\alpha_j - 1\}$; from that the following proposition follows.

Proposition 4.2. *Let S_1^1 denote the unit circle in $\mathbb{C}$. If $A \in \mathbb{R}^{n\times n}$, the following statements are equivalent:*

$$\begin{aligned} &(i) \text{ For all } Y \in \mathrm{Symm}(n)\ \exists X : A^\top XA - X = Y\\ &(ii)\ \alpha_i\alpha_j \neq 1 \text{ for all } \alpha_i, \alpha_j \in^e c(A)\\ &(iii)\ \operatorname{spec}(A) \cap \operatorname{spec}(A^{-1}) = \emptyset\\ &(iv)\ \operatorname{spec}(A) \cap S = \emptyset. \end{aligned}$$

△

A matrix whose spectrum lies inside the unit disc in $\mathbb{C}$ is called a Schur–Cohn matrix; its characteristic polynomial can be called a Schur–Cohn polynomial (analogous to "Hurwitz"). The criterion for a polynomial to have the Schur–Cohn property is given as an algorithm in Henrici [121, pp. 491–494]. For monic quadratic polynomials $\boldsymbol{p}(s) = s^2 + y_1 s + y_2$, the criterion (necessary and sufficient) is

$$|y_2| < 1 \text{ and } |y_1| < 1 + y_2. \tag{4.3}$$

Another derivation of (4.3) is to use the mapping $f : z \to (1 + z)/(1 - z)$, which maps the unit disc in the s plane to the left half of the z plane. Let $\boldsymbol{q}(z) = \boldsymbol{p}(f(z))$. The numerator of $\boldsymbol{q}(z)$ is a quadratic polynomial; condition (4.3) is sufficient and necessary for it to be a Hurwitz polynomial.

The following stability test is over a century old; for some interesting generalizations, see Dirr and Wimmer [73] and its references.

Theorem 4.1 (Eneström–Kakeya). *Let $h(z) := a_n z^n + \cdots + a_1 z + a_0$ be a real polynomial with*

$$a_n \geq \cdots \geq a_1 \geq a_0 \geq 0 \text{ and } a_n > 0.$$

Then all the zeros of h are in the closed unit disc, and all the zeros lying on the unit circle are simple.

4.2 Discrete-Time Control

Discrete-time bilinear control systems on $\mathbb{R}^n$ are given by difference equations

$$x(t+1) = Ax(t) + \sum_{i=1}^m u_i(t)B_i x(t),\ t \in \mathbb{Z}_+, \text{ or briefly } \overset{+}{x} = F(u)x, \tag{4.4}$$

$$\text{where } F(u(t)) := A + \sum_{i=1}^m u_i(t)B_i;$$

the values $u(t)$ are constrained to some closed set Ω; the default is $\Omega = \mathbb{R}^m$, and another common case is $\Omega = [-\mu, \mu]^m$. The appropriate state space here is still $\mathbb{R}^n_*$, but as in Section 4.1 if $A(u)$ becomes singular some trajectories may reach $\{0\}$ in finite time, so the origin may be the terminus of a trajectory.

The transition matrices and trajectories for (4.4) are easily found. The trajectory is a sequence $\{\xi, x(1), x(2), \dots\}$ of states in $\mathbb{R}^n_*$ that are given by a discrete-time transition matrix: $x(t) = P(t; u)\xi$ where

$$P(0; u) := I,\ P(t; u) := F(u(t)) \cdots F(u(1)),\ t \in \mathbb{Z}_+. \tag{4.5}$$

$$\mathbf{S}_\Omega := \{P(t; u) | \, t \in \mathbb{Z}_+, u(\cdot) \in \Omega\} \tag{4.6}$$

is the semigroup of transition matrices of (4.4). When $\Omega = \mathbb{R}^m$, the subscript with $\mathbf{S}$ is suppressed. The transition matrix can be expressed as a series of multilinear terms; if $\overset{+}{X} = AX + uBX,\ X(0) = I$, we have (since $A^0 = I$)

$$P(t; u) = A^t + \sum_{i=1}^{t} A^{t-i} B A^{i-1} u(i) + \cdots + B^t \prod_{i=1}^{t} u(i). \tag{4.7}$$

The set of states $\mathbf{S}\xi$ is called the orbit of $\mathbf{S}$ through ξ. The following properties are easy to check.

1. Conjugacy: if $x = Ty$ then $\overset{+}{y} = TF(u)T^{-1}y$.
2. If for some value μ of the control $\det(F(\mu)) = 0$ then there exist states ξ and finite control histories $\{u\}_1^T$ such that $P(T; u)\xi = 0$. If $F(\mu)$ is nilpotent, trajectories from all states terminate at 0.
3. The location and density of the orbit through ξ may depend discontinuously on A and B (Example 4.1) and its cardinality may be finite (Example 4.2).

Example 4.1 (Sparse orbits). Take $\overset{+}{x} = uBx$, where

$$B := \begin{bmatrix} \cos(\lambda\pi) & -\sin(\lambda\pi) \\ \sin(\lambda\pi) & \cos(\lambda\pi) \end{bmatrix}.$$

For $\lambda = p/q$, p and q integers, $\mathbf{S}\xi$ consists of q rays; for λ irrational, $\mathbf{S}\xi$ is a dense set of rays. △

Example 4.2 (Finite group actions). A discrete-time switched system acting on $\mathbb{R}^n$ has a control rule u which at $t \in \mathbb{Z}_+$ selects a matrix from a list $\mathcal{F} := \{A(1), \dots, A(m)\}$ of $n \times n$ matrices, and obeys

$$x(t+1) = A(u_t)x(t),\ x(0) = \xi \in \mathbb{R}^n,\ u_t \in 1 \mathinner{\ldotp\ldotp} m.$$

$$x(t) = P(t; u)\xi \text{ where } P(t; u) = \prod_{k=1}^{t} A(u_k)$$

is the transition matrix. The set of all such matrices is a matrix semigroup. If $A(u) \in \mathcal{F}$ then $A^{-1}(u) \in \mathcal{F}$ and this semigroup is the group $\mathbf{G}$ generated by $\mathcal{F}$; $\mathbf{G}\xi$ has cardinality no greater than that of $\mathbf{G}$. If the entries of each A_i are integers and $\det(A_i) = 1$ then $\mathbf{G}$ is a finite group. △

4.3 Stabilization by Constant Inputs

Echoing Chapter 3 in the discrete-time context, one may pose the problem of stabilization with constant control, which here is a matter of finding μ such that $A + \mu B$ is Schur–Cohn. The $n = 2$, $m = 1$ case again merits some discussion. Here, as in Section 3.2.1, the characteristic polynomial $\boldsymbol{p}_{A+\mu B}(s) := s^2 + sy_1(\mu) + y_2(\mu)$ can be checked by finding any μ interval on which it satisfies the $n = 2$ Schur–Cohn criterion

$$|y_2(\mu)| < 1, \ |y_1(\mu)| < 1 + y_2(\mu).$$

Graphical methods as in Fig. 3.1 can be used to check the inequalities.

Example 4.3. For $n = 2$, suppose $\mathrm{tr}(A) = 0 = \mathrm{tr}(B)$. We see in this case that $\boldsymbol{p}_{A+\mu B} = s^2 + \det(A + \mu B)$ is a Schur–Cohn polynomial for all values of μ such that

$$\left|\det(A) - \mu\,\mathrm{tr}(AB) + \mu^2 \det(B)\right| < 1. \qquad \triangle$$

Problem 4.1. For discrete-time bilinear systems on $\mathbb{R}^2_*$ conjecture and prove some partial analog of Theorem 3.1, again using the three real canonical forms of B. Example 4.3 is pertinent.

If we take $F_\mu = A + \mu B$, then the discrete-time Lyapunov direct method can be used to check F for convergence given a range of values of μ. The criterion is as follows (see (A.9)).

Proposition 4.3. *The spectrum of matrix F lies inside the unit disc if and only if there exist symmetric matrices Q, P such that*

$$Q \gg 0, \ P \gg 0, \ F_\mu^\top Q F_\mu - Q = -P, \tag{4.8}$$
$$\text{or} \quad (F_\mu^\top \otimes F_\mu)Q^\flat - Q^\flat = -P^\flat. \qquad \triangle$$

Although some stabilization methods in Chapter 3, such as Proposition 3.2, have no analog for discrete-time systems, there are a few such as the triangularization method for which the results are parallel. Let $\{\alpha_1, \dots, \alpha_n\}$ and $\{\beta_1 \dots \beta_n\}$ again be the eigenvalues of A and B, respectively. If A and B are simultaneously similar to (possibly complex) upper triangular matrices A_1, B_1 then the diagonal elements of $A_1 + \mu B_1$ are $\alpha_i + \mu\beta_i$; find all μ such that

$$|\alpha_i + \mu\beta_i| < 1, \ i \in 1, \dots, n.$$

The simultaneous triangularizability of A, B, we have already observed, is equivalent to the solvability of the Lie algebra $\mathfrak{g} = \{A, B\}_{\mathcal{L}}$; one might check it with the Cartan–Killing form χ defined on $(\mathfrak{g}, \mathfrak{g})$ as in Section B.2.1.

4.4 Controllability

The definitions in Section 1.6.2 can be translated into the language of discrete-time control systems $\overset{+}{x} = f(x, u)$, where $f : \mathbb{R}^n \times \mathbb{R}^m$ takes finite values; the most important change is to replace $t \in \mathbb{R}_+$ with $k \in \mathbb{Z}_+$. Without further assumptions on f, trajectories exist and are unique. In computation, one must beware of systems whose trajectories grow superexponentially, like $\overset{+}{x} = x^2$.

Definitions 1.9 and 1.10 and the Continuation Lemma (Lemma 1.1, page 22) all remain valid for discrete-time systems.

Corresponding to the symmetric systems (2.2) of Chapter 2 in the discrete-time case are their Euler discretizations with step-size τ, which are of the form

$$\overset{+}{x} = \left(I + \tau \sum_{i=1}^{m} u_i(k) B_i\right) x,\ x(0) = \xi. \tag{4.9}$$

For any matrix B and for both positive and negative τ

$$\lim_{k->\infty} \left\| \left(I + \frac{\tau}{k} B\right)^k - e^{\tau B} \right\| = 0,$$

so matrices of the form $\exp(\tau B)$ can be approximated in norm by matrices in $\mathbf{S}$, the semigroup of transition matrices of (4.9). Therefore, $\mathbf{S}$ is dense in the Lie group $\mathbf{G}$ that (2.2) provides. The inverse of a matrix $X \in \mathbf{S}$ is not guaranteed to lie in $\mathbf{S}$. The only case at hand where it does so is nearly trivial.

Example 4.4. If $A^2 = 0 = AB = B^2$ then $(I + uA + vB)^{-1} = I - uA - vB$. △

4.4.1 Invariant Sets

A polynomial p for which $p(Ax) = p(x)$ is called an invariant polynomial of the semigroup $\{A^j | j \in \mathbb{Z}_+\}$. If $1 \in \operatorname{spec}(A)$, there exists a linear invariant polynomial $c^{\top} x$. If there exists $Q \in \operatorname{Symm}(n)$ (with any signature) such that $A^{\top} Q A = Q$, then $x^{\top} Q x$ is invariant for this semigroup. Relatively invariant polynomials p such that $p(Ax) = \lambda p(x)$ can also be defined.

In the continuous-time case, we were able to locate the existence of invariant affine varieties for bilinear control systems. The relative ease of the computations in that case is because the varieties contain orbits of matrix Lie groups.[1] It is easy to find examples of invariant sets, algebraic or not; see those in Examples 4.1 and 4.2. Occasionally uncontrollability of $\overset{+}{x} = Ax + uBx$ can be shown from the existence of one or more linear functions $c^\top x$ that are invariant for any u and thus satisfy both $c^\top A = c^\top$ and $c^\top B = c^\top$. Another invariance test, given in Cheng [57], follows.

Proposition 4.4. *If there exists a nonzero matrix C that commutes with A and B and has a real eigenvalue λ, then the eigenspace $L_\lambda := \{x | Cx = \lambda x\}$ is invariant for* (4.4).

Proof. Choose any $x(0) = \xi \in L$. If $x(t-1) \in L$, then

$$Cx(t) = C(A + u(t-1)B)x(t-1) = (A + u(t-1)B)\lambda x(t-1) = \lambda x(t);$$

by induction all trajectories issuing from ξ remain in L_λ. □

If for some fixed k there is a matrix $C \neq 0$ that, although it does not commute with A and B, does commute for some $t > 1$ with the transition matrix $P(t;u)$ of (4.4) for all $u(0), \dots, u(t-1)$, the union of its eigenvectors L_λ is invariant. The following example illustrates this.

Example 4.5. For initial state $\xi \in \mathbb{R}^2$, suppose

$$\overset{+}{x} = \begin{bmatrix} 0 & 1+u \\ -1 & 0 \end{bmatrix} x; \;\; P(2;u) = \begin{bmatrix} -1-u(1) & 0 \\ 0 & -1-u(0) \end{bmatrix}. \text{ Let } C := \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}$$

with real $\alpha \neq \beta$. C commutes with $P(2;u)$ but not with $P(1;u)$, and the invariant set $L := \{x | \, x_1 x_2 = 0\}$ is the union of the eigenspaces L_α and L_β, neither of which is itself invariant. However, if $\xi_1\xi_2 \neq 0$ and u is unconstrained then for any $\zeta \in \mathbb{R}^2$, $x(2) = \zeta$ is satisfied if $u(0) = -1 - \zeta_1/\xi_1$, $u(1) = -1 - \zeta_2/\xi_2$. For $\xi \notin L$, $\mathcal{A}(\xi) = \mathbb{R}^2$ but for certain controls there is a terminal time k_0 such that $P(t;u)\xi \in L$ for $k \geq k_0$. △

4.4.1.1 Traps

Finding traps for (4.4) with unrestricted controls is far different from the situation in Section 3.3.2. With $\Omega = \mathbb{R}^m$, given any cone K other than a subspace or half-space, K cannot be invariant: for $x \in K$ part of the set $\{Ax + \sum u_i B_i x | \, u \in \mathbb{R}^m\}$ lies outside K. With compact Ω, the positive orthant is invariant if and only if every component of $F(u)$ is nonnegative for all $u \in \Omega$.

[1] Real algebraic geometry is applied to the accessibility properties of $\overset{+}{x} = (A + uB)x$ and its polynomial generalizations by Sontag [244, 247].

4.4.2 Rank-One Controllers

Some early work on bilinear control made use of a property of rank-one matrices to transform some special bilinear systems to linear control systems. If an $n \times n$ matrix B has rank 1, it can be factored in a family of ways $B = bc^\top = bSS^{-1}c^\top$ as the outer product of two vectors, a so-called dyadic product. In Example 4.5, for instance, one can write $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix}$. Example 4.5 and Exercise 1.13 will help in working through what follows.

The linear control system

$$\overset{+}{x} = Ax + bv,\ y = c^\top x;\ \ x \in \mathbb{R}^n, v(t) \in \mathbb{R}; \tag{4.10}$$

with multiplicative feedback control $v(t) = u(t)c^\top x(t)$, becomes a bilinear system $\overset{+}{x} = Ax + u(bc^\top)x$. If (4.10) is controllable, necessarily the controllability matrix $\mathbf{R}(A; b)$ defined in (1.43) has rank n and given $\zeta \in \mathbb{R}^n$, one can find a control sequence $\{v(0), \ldots, v(n-1)\}$ connecting ξ with ζ by solving

$$\zeta - A^n\xi = \mathbf{R}(A; b)\mathbf{v} \text{ where } \mathbf{v} := \mathrm{col}(v(n-1), v(n-2), \ldots, v(0)).$$

Consider the open half-spaces $U_+ := \{x|\, c^\top x > 0\}$ and $U_- := \{x|\, c^\top x < 0\}$. It was suggested in Goka et al. [105][2] that the bilinear system

$$\overset{+}{x} = Ax + ubc^\top x,\ \ x(0) = \xi,\ \ u(t) \in \mathbb{R}; \tag{4.11}$$

$$\text{with}\ \ u(t) = \begin{cases} v(t)/(c^\top x(t)), & c^\top x(t) \neq 0, \\ \text{undefined}, & c^\top x = 0 \end{cases} \tag{4.12}$$

has the same trajectories as (4.10) on the submanifold $U_- \bigcup U_+ \subset \mathbb{R}^n$. This idea, under such names as division controller has had an appeal to control engineers because of the ease of constructing useful trajectories for (4.10) and using (4.12) to convert them to controls for (4.11).

However, this trajectory equivalence fails if $x(t) \in L := \{x|\, c^\top x = 0\}$ because $\overset{+}{x} = Ax$ on L and $u(t)$ has no effect. The nullspace of

$$\mathbf{O}(c^\top; A) := \mathrm{col}(c^\top, c^\top A, \ldots, c^\top A^{n-1})$$

is a subset of L and is invariant for $\overset{+}{x} = Ax$. For this reason, let us also assume that the rank of $\mathbf{O}(c^\top; A)$ is full.

Given a target state ζ and full rank of $\mathbf{R}(A; b)$, there exists an n-step control $\mathbf{v}$ for (4.10) whose trajectory, the *primary trajectory*, satisfies $x(0) = \xi$ and $x(n) = \zeta$. For generic choices of ξ and ζ, the feedback control $u(t) := v(t)/(c^\top x(t))$ will provide a control for (4.11) that provides the same trajectory connecting ξ to ζ. However, for other choices some state $x(t)$ on the primary trajectory

[2] Comparing [87, 105] our $B = ch^\top$ is their $B = bc^\top$.

is in the hyperplane L; the next control $u(t)$ has no effect and the trajectory fails to reach the target. A new control for (4.11) is needed; to show one exists, two more generic[3] hypotheses are employed in [105, Theorem 3] and its corollary, whose proofs take advantage of the flexibility obtained by using control sequences of length $2n$. It uses the existence of A^{-1}, but that is not important; one may replace A with $A + \mu B$, whose nonsingularity for many values of μ follows from the full rank of $\mathbf{R}(A; b)$ and $\mathbf{O}(c^{\top}; A)$.

Theorem 4.2. *Under the hypotheses* $\operatorname{rank} \mathbf{R}(A; b) = n$, $\operatorname{rank} \mathbf{O}(c^{\top}; A) = n$, $\det(A) \neq 0$, $c^{\top} b \neq 0$ *and* $c^{\top} A^{-1} b \neq 0$, *the bilinear system* (4.11) *is controllable.*

Proof. It can be assumed that $\xi \notin L$ because using repeated steps with $v = 0$, under our assumptions we can reach a new initial state ξ' outside L, and that state will be taken as ξ. Using the full rank of $\mathbf{O}(c^{\top}; A)$, if $\zeta \in L$ one may find the first j between 0 and n such that $c^{\top} A^{-j} \zeta \neq 0$, and replace ζ with $\zeta' := A^{-k} \zeta$. From ζ', the original ζ can be reached in k steps of (4.10) with $v = 0$. Thus all that needs to be shown is that primary trajectories can be found that connect ξ with ζ and avoid L.

The linear system (4.10) can be looked at backward in time,[4]

$$x(t) = A^{-1} x_{k+1} - v(t) A^{-1} b, \; x(T) = \zeta.$$

The assumption that $c^{\top} A^{-1} b \neq 0$ means that at each step k the value of $v(t)$ can be adjusted slightly to make sure that $c^{\top} x(t) \neq 0$. After n steps, the trajectory arrives at a state $y = x(T - n) \notin L$. In forward time, ζ can be reached from y via the n-step trajectory with control $v'(t) = v(n - k)$ that misses L; furthermore, there is an open neighborhood $N_{\epsilon}(y)$ in which each state has that same property.

There exists a control v'' and corresponding trajectory that connect ξ to y in n steps. Using the assumption $c^{\top} b \neq 0$ in (4.10), small adjustments to the controls may be used to ensure that each state in the trajectory misses L and the endpoint $x(n)$ remains in $N_{\epsilon}(y)$. Accordingly, there are a control v^{*} and corresponding trajectory from $x(n)$ to ζ that misses L. The concatenation $v^{*} \star v''$ provides the desired trajectory from ξ to ζ. □

Example 4.6 (Goka [105]). For dimension two, (4.11) can be controllable on $\mathbb{R}^2_*$ even if $c^{\top} b = 0$, provided that $\mathbf{O}(c^{\top}; A)$ and $\mathbf{R}(A; b)$ both have rank 2. Define the scalars

---

[3] Saying that the conditions are generic means that they are true unless the entries of A, b, c satisfy some finite set of polynomial equations, which in this case are given by the vanishing of determinants and linear functions. Being generic is equivalent to openness in the Zariski topology (Section C.2.3) on parameter space.

[4] This is not at all the same as running (4.11) backward in time.

$$\alpha_1 = c^\top \xi,\ \alpha_2 = c^\top A\xi;\ \beta_1 = c^\top b,\ \beta_2 = c^\top Ab;\ \text{then } x(1) = A\xi + u(0)\alpha_1 b,$$
$$\zeta = A^2\xi + u(0)\alpha_1 Ab + \alpha_2 u(1)b + u(0)u(1)\beta_1\alpha_1 b.$$
$$\text{If } \beta_1 = 0,\ \begin{bmatrix} u(1) \\ u(0) \end{bmatrix} = \begin{bmatrix} \alpha_2 b & \alpha_1 Ab \end{bmatrix}^{-1} (\zeta - A^2\xi).$$

The full rank of $\mathbf{O}(c^\top; A)$ guarantees that α_1, α_2 are nonzero, so ζ can be reached at $t = 2$. A particularly simple example is the system $\overset{+}{x}_1 = (u-1)x_2, \overset{+}{x}_2 = x_1$. △

For (4.11) with $B = c^\top b$, Evans and Murthy [87] improved Theorem 4.2 by dealing directly with (4.11) to avoid all use of (4.10).

Theorem 4.3. *In* (4.11), *let*

$$\mathfrak{I} := \left\{ i \middle| c^\top A^{i-1} b \neq 0,\ 0 < i < n^2 \right\}$$

and let K be the greatest common divisor of the integers in $\mathfrak{I}$. Then (4.11) *is controllable if and only if*

(1) $\operatorname{rank} \mathbf{R}(A; b) = n = \operatorname{rank} \mathbf{O}(c^\top; A)$ *and*

(2) $K = 1$.

Hypothesis (2) is interesting: if $K \geq 2$ then the subspace spanned by $\mathbf{R}(A^k; b)$ is invariant and, as in Example 4.5, may be part of a cycle of subspaces. The proof in [87] is similar to that of Theorem 4.2; after using the transformation $x(t) = A^k z(t)$, the resulting time-variant bilinear system in z and its time-reversal (which is rational in u) are used to find a ball in $\mathbb{R}^n_*$ that is attainable from ξ and from which a target ζ can be reached. Extending the (perhaps more intuitive) method of proof of Theorem 4.2 to obtain Theorem 4.3 seems feasible. Whether either approach can be generalized to permit multiple inputs (with each B_i of rank one) seems to be an open question. △

4.4.3 Small Controls: Necessity

Controllability of the single-input constrained system

$$\overset{+}{x} = Ax + Bxu,\ |u(t)| \leq \mu \tag{4.13}$$

using indefinitely small μ was investigated in the dissertations of Goka [274] and Cheng [57]; compare Section 3.5. As in Chapter 3, (4.13) will be called small-controllable if for every bound $\mu > 0$ it is controllable on $\mathbb{R}^n_*$.

The following two lemmas are much like Lemma 3.6, and are basic to the stability theory of difference equations, so worth proving; their proofs are almost the same.[5]

[5] Cheng [56] has a longer chain of lemmas.

Lemma 4.1. *If A has any eigenvalue α_p such that $|\alpha_p| > 1$, then there exists $\rho > 1$ and real $R = R^\top$ such that $A^\top RA = \rho^2 R - I$ and $V(x) := x^\top Rx$ has a negative domain.*

Proof. For some vector $\zeta = x + y\sqrt{-1} \neq 0$, $A\zeta = \alpha_p\zeta$. Choose any ρ for which $1 < \rho < |\alpha_p|$ and $\rho \neq |\alpha_i|$ for all the eigenvalues $\alpha_i \in \operatorname{spec}(A)$. Let $C := \rho^{-1}A$, then $\operatorname{spec}(C) \cap S_1^1 = \emptyset$ by Proposition 4.2. Consequently $C^\top RC - R = -\rho^{-2}I$ has a unique symmetric solution R.

$$A^\top RA - \rho^2 R = -I\rho^2\zeta^* R\zeta - \zeta^*\zeta = \zeta^* A^\top RA\zeta \text{ and}$$
$$\zeta^* A^\top RA\zeta = \alpha_p^*\alpha_p\zeta^* R\zeta. \text{ Since } \rho^2 < |\alpha_p|^2,$$
$$\zeta^* R\zeta = \frac{|\zeta|^2}{\rho^2 - |\alpha_p|^2} < 0. \tag{4.14}$$

Since $\zeta^* R\zeta = x^\top Rx + y^\top Ry$, the inequality (4.14) becomes

$$0 > x^\top Rx + y^\top Ry = V(x) + V(y),$$

so V has negative domain, to which at least one of x, y belongs. □

Lemma 4.2. *If A has any eigenvalue α_q such that $|\alpha_q| < 1$, then there exists real σ and matrix $Q = Q^\top$ such that $A^\top QA - \sigma^2 Q = -I$ and $V(x) := x^\top Qx$ has a negative domain.*

Proof. Let $\zeta = x + y\sqrt{-1} \neq 0$ satisfy $A\zeta = \alpha_q\zeta$. Choose a σ that satisfies $\sigma \neq |\alpha_i|$ for all the eigenvalues $\alpha_i \in \operatorname{spec}(A)$ and $1 > |\alpha_q| > \sigma$. There exists a unique symmetric matrix Q such that $A^\top QA = \sigma^2 Q - I$. Again evaluating $\zeta^* A^\top QA\zeta$, and using $\alpha_q^*\alpha_q > \sigma^2$,

$$\sigma^2\zeta^* Q\zeta - \zeta^*\zeta = \alpha_q^*\alpha_q\zeta^* Q\zeta;$$
$$\zeta^* Q\zeta = \frac{|\zeta|^2}{\sigma^2 - |\alpha_q|^2} < 0.$$

Again either the real part x or imaginary part y of ζ must belong to the negative domain of V. □

The following is stated in [56, 57] and proved in [56]; the proof here, inclusive of the lemmas, is shorter and maybe clearer. The scheme of the indirect proof is that if there are eigenvalues off the unit circle there are gauge functions that provide traps, for sufficiently small control bounds.

Proposition 4.5. *If* (4.13) *is small-controllable then the eigenvalues of A are all on the unit circle.*

Proof. If not all the eigenvalues of A lie on S_1^1, there are two possibilities.

First suppose A has $A\zeta = \alpha_p\zeta$ with $|\alpha_p| > 1$. With the matrix R from Lemma 4.1 take $V_1(x) := x^\top Rx$ and look at its successor

$$V_1(\overset{+}{x}) = x^\top(A+uB)^\top R(A+uB)x = x^\top A^\top RAx + W_1(x,u) \text{ where}$$
$$W_1(x,u) := ux^\top(B^\top RA + A^\top RB + u^2B^\top RB)x.$$

Using Lemma 4.1, for all x

$$V_1(\overset{+}{x}) = \rho^2 x^\top Rx - x^\top x + W_1(x,u);$$
$$V_1(\overset{+}{x}) - V_1(x) = (\rho^2-1)x^\top Rx - x^\top x + W_1(x,u)$$
$$= (\rho^2-1)V_1(x) - x^\top x + W_1(x,u).$$

There exists an initial condition ξ such that $V_1(\xi) < 0$, from the lemma. Then

$$V_1(\overset{+}{\xi}) - V_1(\xi) = (\rho^2-1)V_1(\xi) - \xi^\top\xi + W_1(\xi,u) \text{ and since}$$
$$(\rho^2-1)V_1(\xi) < -(\rho^2-1)\delta\|\xi\|^2,$$
$$V_1(\overset{+}{\xi}) - V_1(\xi) < -\|\xi\|^2 + W_1(\xi,u).$$

$W_1(\xi,u) < \mu\beta_1\|\xi\|^2$ where β_1 depends on A, B, and ρ through (4.4.3). Choose $|u| < \mu$ sufficiently small that $\mu\beta < 1$; then V_1 remains negative along all trajectories from ξ, so the system cannot be controllable, a contradiction.

Second, suppose A has $A\zeta = \alpha_q\zeta$ with $|\alpha_q| < 1$. From Lemma 4.2, we can find $\sigma > 0$ such that $1 > |\alpha_q| > \sigma$ and $A^\top QA - \sigma^2 Q = I$ has a unique solution; set $V_2(x) := x^\top Qx$. As before, get an estimate for

$$W_2(x,u) := ux^\top(B^\top QA + A^\top QB + u^2B^\top QB)x : \ W_2(\xi,u) < \mu\beta_2\|\xi\|^2.$$
$$\exists\xi \text{ s.t. } V_2(\xi) < 0 \text{ by Lemma 4.2;}$$
$$V_2(\overset{+}{\xi}) - V_2(\xi) = (\sigma^2-1)x^\top Qx - x^\top x + W_2(x,u)$$
$$= (\sigma^2-1)V_2(\xi) - \xi^\top\xi + W_2(\xi,u)$$
$$= (\sigma^2-1)V_2(\xi) - \|\xi\|^2 + W_2(\xi,u) < 0,$$

if the bound μ on u is sufficiently small, providing the contradiction. □

4.4.4 Small Control: Sufficiency

To obtain sufficient conditions for small-controllability, note first that (as one would expect after reading Chapter 3) the orthogonality condition on A can be weakened to the existence of $Q \gg 0$, real μ and integer r such that $C^\top QC = Q$ where $C := (A+\mu B)^r$.

Secondly, a rank condition on (A, B) is needed. This parallels the *ad*-condition on A, B and neutrality condition in Section 3.4.1 (see Lemma 3.3). Referring to (4.5), the Jacobian matrix of $P(k;u)x$ with respect to the history

$\mathfrak{U}_k = (u(0), \ldots, u(k-1)) \in \mathbb{R}^k$ is

$$J_k(x) := \frac{\partial}{\partial \mathfrak{U}_k} ((A + u(t-1)B) \cdots (A + u(0)B)x)_{\mathfrak{U}_k=0}$$
$$= \left[A^k Bx \cdots A^{k-i}BA^i x \cdots BA^k x\right]. \tag{4.15}$$
$$\text{If } \exists \kappa : \ \operatorname{rank} J_\kappa(x) = n \text{ on } \mathbb{R}^n_*, \tag{4.16}$$

then (4.5) is said to satisfy the local rank condition. To verify this condition, it suffices to show that the origin is the only common zero of the n-minors of $J_k(x)$ for $k = n^2-1$ (at most). Calling (4.16) an *ad*-condition would be somewhat misleading, although its uses are similar. The example in Section 4.5 will show that full local rank is not a necessary condition for controllability on $\mathbb{R}^2_*$.

Proposition 4.6. *If there exists $Q \gg 0$ and $\kappa \in \mathbb{N}$ s.t. $(A^\kappa)^\top Q A^\kappa = Q$ and if the local rank condition is met, then* (4.13) *is small-controllable.*

Proof. For any history $\mathfrak{U}_\kappa$ for which the $u(i)$ are sufficiently close to 0, the local rank condition implies that the polynomial map $\mathbf{P}_\kappa : U \to P(\kappa; u)x$ is a local diffeomorphism. That A^κ is an orthogonal matrix with respect to Q ensures that I is in the closure of $\{A^k | \, k \in \mathbb{N}\}$. One can then conclude that any initial x is in its own attainable set and use the Continuation Lemma 1.1. □

Example 4.7. Applying Proposition 4.6 to $\overset{+}{x} = Ax + uBx$ with

$$A := \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \ B := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ J_1(x) = \frac{1}{\sqrt{2}} \begin{bmatrix} x_2 - x_1 & x_1 + x_2 \\ x_1 + x_2 & x_1 - x_2 \end{bmatrix},$$

$A^\top A = I$ and $\det(J_1(x)) = -(x_1^2 + x_2^2)$, it is small-controllable. △

4.4.5 Stabilizing Feedbacks

The local rank condition and Lyapunov functions can be used to get quadratic stabilizers in a way comparable to Theorem 3.5 — or better, with feedbacks of class $\mathcal{H}^2_0$ (compare Theorem 3.8) for small-controllable systems.

Theorem 4.4. *Suppose that the system $\overset{+}{x} = Ax + uBx$ on $\mathbb{R}^n_*$ with unrestricted controls satisfies*

(a) there exists $Q \gg 0$ for which $A^\top QA = Q$;

(b) B is nonsingular; and

(c) the local rank condition (4.16) *is satisfied. Then there exists a stabilizing feedback $\phi \in \mathcal{H}^2_0$.*

Proof. Use the positive definite gauge function

$$V(x) := x^\top Qx \text{ and let } W(x,u) := V(\overset{+}{x}) - V(x).$$

$$W(x,u) = 2u\Big(x^\top B^\top QAx\Big) + u^2 x^\top B^\top QBx;$$

$$W_0(x) := \min_u W(x,u) = -\frac{(x^\top B^\top QAx)^2}{x^\top B^\top QBx} \leq 0 \;\text{ is attained for}$$

$$\phi(x) = \arg\min_u(W(x,u)) = -\frac{x^\top B^\top QAx}{x^\top B^\top QBx}.$$

The trajectories of $\dot{x} = Ax + \phi(x)Bx$ will approach the quadric hypersurface $\mathcal{E}$ where $W_0(x)$ vanishes; from the local rank condition, $\mathcal{E}$ has no invariant set other than $\{0\}$. □

4.4.6 Discrete-Time Biaffine Systems

Discrete-time biaffine systems, important in applications,

$$\overset{+}{x} = Ax + a + \sum_{i=1}^{m} u_i(B_i x + b_i),\; a, b \in \mathbb{R}^n,\; y = Cx \in \mathbb{R}^p \tag{4.17}$$

can be represented as a homogeneous bilinear systems on $\mathbb{R}^{n+1}$ by the method of Section 3.8.1:

$$z := \begin{bmatrix} x \\ x_{n+1} \end{bmatrix};\; \overset{+}{z} = \begin{bmatrix} A & a \\ \mathbf{0}_{1,n} & 0 \end{bmatrix} z + \sum_{i=1}^{m} u_i \begin{bmatrix} B_i & b_i \\ \mathbf{0}_{1,n} & 0 \end{bmatrix} z;\; y = (C,0)z.$$

Problem 4.2. In the small-controllable Example 4.7, $A^8 = I$, so by (1.15) $\Lambda_d(A;8) = 0$. Consider the biaffine system $\overset{+}{x} = Ax + uBx + a$, where $\|a\| \neq 0$. With $u = 0$ as control $x(8) = \xi$. Find additional hypotheses under which the biaffine system is controllable on $\mathbb{R}^n$. Are there discrete-time analogs to Theorems 3.13 and 3.14?

4.5 A Cautionary Tale

This section has appeared in different form as [84].[6]

One might think that if the Euler discretization (4.9) of a bilinear system is controllable, so is the original system. The best-known counterexample to such a hope is trivial: $\overset{+}{x}_1 = x_1 + ux_1$, the Euler discretization of $\dot{x}_1 = ux_1$.

For single-input systems, there are more interesting counterexamples that come from continuous-time uncontrollable single-input symmetric systems on $\mathbb{R}^2_*$. Let

$$\dot{x} = u(t)Bx, \text{ where } B := \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix}, \ x(0) = \xi \tag{4.18}$$

$$\text{with trajectory } \gamma_\xi := \left\{ e^{B\int_0^t u(s)\,ds}\xi, \ t \geq 0 \right\}$$

whose attainable set from each ξ is a logarithmic spiral curve. The discrete control system obtained from (4.18) by Euler's method is

$$\overset{+}{x} = (I + v(t)B)x; \ x(0) = \xi; \ v(t) \in \mathbb{R}. \tag{4.19}$$

Assume that λ is negative. For control values $v(t) = \omega$ in the interval

$$U := \frac{2\lambda}{\lambda^2 + \mu^2} < \omega < 0, \quad \|I + \omega B\|^2 = 1 + 2\omega\lambda + \omega^2(\lambda^2 + \mu^2) < 1.$$

With $v(t) \in U$, the trajectory of (4.19) approaches the origin, and for a constant control $v(t) = \omega \in U$ that trajectory is a discrete subset of the continuous spiral

$$\eta_\xi := \{(1 + \omega B)^t \xi \,|\, t \in \mathbb{R}\},$$

exemplifying Proposition 4.1.

A conjecture, *If* $\lambda < 0$ *then* (4.19) *is controllable,* was suggested by the observation that for any $\xi \in \mathbb{R}^2_*$ the union of the lines tangent to each spiral η_ξ covers $\mathbb{R}^2_*$ (as do the tangents to any asymptotically stable spiral trajectory). It is not necessary to prove the conjecture, since by using some symbolic calculation one can generate families of examples like the one that follows.

Proposition 4.7. *There exists an instance of B such that* (4.19) *is controllable.*

Proof. Let $\bar{v} = 2\lambda/(\lambda^2 + \mu^2)$; then the trajectory $X(\bar{v}, k)\xi$, $k \in \mathbb{Z}$ lies on a circle of radius $\|\xi\|$. To show controllability, we can use any B for which there are nominal trajectories using $v(t) = \bar{v}$ that take ξ to itself at some time $k = N - 1$; then change the value $v(N-1)$ and add a new control $v(N)$ to reach nearby

[6] Erratum: in [84, Eq. 5], the value of μ in [84, Eq. 5] should be $\sqrt{3}/2$ there and in the remaining computations. This example has been used in my lectures since 1974, but the details given here are new.

states noting that if $v(N) = 0$ the trajectory still ends at ξ. The convenient case $N = 6$ can be obtained by choosing

$$B = \boldsymbol{\alpha}(\lambda, \mu),\ \lambda := \frac{-1}{2},\ \mu := \frac{-\sqrt{3}}{2}.\ \text{ For } \bar{v} := 1,\ (I + \bar{v}B)^6 = I. \tag{4.20}$$

The two-step transition matrix for (4.19) is $f(s,t) := (I + tB)(I + sB)$. For $x \in \mathbb{R}^2_*$, the Jacobian determinant of the mapping $(s,t) \mapsto f(s,t)x$ is, for matrices $B = \lambda I + \mu J$,

$$\frac{\partial(f(s,t)x)}{\partial(s,t)} = \mu(\lambda^2 + \mu^2)(t - s)\|x\|^2 = \frac{\sqrt{3}}{2}(t - s)\|x\|^2$$

for our chosen B.

Let $s = v(5) = \bar{v}$ and add the new step using the nominal value $v(6) = t = 0$, to make sure that ξ is a point where this determinant is nonzero. Then (invoking the implicit function theorem) we will construct controls that start from ξ and reach any point in an open neighborhood $U(\xi)$ of ξ that will be constructed explicitly as follows.

Let $v(j) := \bar{v}$ for the five values $j \in 0, \ldots, 4$. For target states $\xi + x$ in $U(\xi)$, a changed value $v(5) := s$ and a new value $v(6) := t$ are needed; finding them is simplified, by the use of (4.20), to the solution of two quadratic equations in (s,t)

$$(I + tB)(I + sB)\xi = (I + \bar{v}B)(\xi + x).\ \text{ If } \xi = \begin{bmatrix}1\\0\end{bmatrix}$$

$$\text{then in either order } (s,t) = \frac{1}{6}\left(3 + 2\sqrt{3}x_2 \pm \sqrt{3}\sqrt{3 + 12x_1 + 4x_2^2}\right),$$

which are real in the region $U(\xi) = \{\xi + x \mid (3 + 12x_1 + 4x_2^2) > 0\}$ that was to be constructed. If $P \in \boldsymbol{\alpha}(\mathbb{C}_*)$ then

$$PB = BP \text{ so } Pf(x,t)\xi = f(s,t)P\xi = Px$$

generalizes our construction to any ξ because $\boldsymbol{\alpha}(\mathbb{C}_*)$ is transitive on $\mathbb{R}^2_*$. Apply the Continuation Lemma 1.1 to conclude controllability. □

If the spiral degenerates ($B = J$) to a circle through ξ, the region exterior to that circle (but not the circle itself) is easily seen to be attainable using the lines $\{(I + tJ)x \mid t \in \mathbb{R}\}$, so strong accessibility can be concluded, but not controllability.

Finally, since $A = I$ the local rank is $\operatorname{rank}[Bx] = 1$, but the multilinear mapping $\mathbf{P}_6$ is invertible. So full local rank is *not* necessary for controllability of (4.13) on $\mathbb{R}^2_*$. Establishing whether the range of $\mathbf{P}_k$ is open in $\mathbb{R}^n_*$ for some k is a much more difficult algebraic problem than the analogous task in Chapter 2.

Chapter 5
Systems with Outputs

To begin, recall that a time-invariant linear dynamical system on $\mathbb{R}^n$ with a linear output mapping

$$\dot{x} = Ax, \; y = Cx \in \mathbb{R}^p \tag{5.1}$$

is called observable if the mapping from initial state ξ to output history $\mathfrak{Y}_T := \{y(t),\ t \in [0, T]\}$ is one-to-one.[1] The Kalman observability criterion (see Proposition 5.1) is that (5.1) is observable if and only if

$$\operatorname{rank} \mathbf{O}(C; A) = n \text{ where } \mathbf{O}(C; A) := \operatorname{col}(C, CA, \ldots, CA^{n-1}) \tag{5.2}$$

is called the observability matrix. If $\operatorname{rank} \mathbf{O}(A; \mu) = n$ then $\{C, A_\mu\}$ is called an observable pair; see Section 5.2.1.

Observability theory for control systems has new definitions, new problems, and especially the possibility of recovering a state space description from the input–output mapping; that is called the realization problem.

Contents of This Chapter

After a discussion in Section 5.1 of some operations that compose new control systems from old, in Section 5.2 state observability is studied for both continuous and discrete-time systems. State observers (Section 5.3) provide a method of asymptotically approximating the system's state ξ from input and output data. The estimation of ξ in the presence of noise is not discussed.

Parameter identification is briefly surveyed in Section 5.4. The realization of a discrete-time input–output mapping as a biaffine or bilinear control system, the subject of much research, is sketched in Section 5.5. As an example, the linear case is given in Section 5.5.1; bilinear and biaffine system realization methods are developed in Section 5.5.2. Section 5.5.3 surveys related work on discrete-time biaffine and bilinear systems.

---

[1] The discussion of observability in Sontag [249] is helpful.

D.L. Elliott, *Bilinear Control Systems*, Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_5, © Springer Science+Business Media B.V. 2009

In Section 5.6 continuous-time Volterra expansions for the input–output mappings of bilinear systems are introduced. In continuous time the realization problem is analytical; it has been addressed either using Volterra series as in Rugh [228] or, as in Isidori [140, Ch. 3], with the Chen–Fliess formal power series for input–output mappings briefly described in Section 8.2. Finally, Section 5.7 is a survey of methods of approximation with bilinear systems.

5.1 Compositions of Systems

To control theorists it is natural to look for ways to compose new systems from old ones given by input–output relations with known initial states. The question here is, for what types of composition and on what state space is the resulting system bilinear? Here only the single-input single-output case will be considered.

We are given two bilinear control systems on $\mathbb{R}^n$ with the same input u; their state vectors are p, q and their outputs are $y_1 = a^\top p$, $y_2 = b^\top q$, respectively. For simplicity of exposition the dimensions of the state spaces for (5.3) and (5.4) are both n. Working out the case where they differ is left as an exercise.

$$\dot{p} = (A_1 + uB_1)p, \;\; y_1(t) = a^\top \Phi_1(t;u)p(0), \tag{5.3}$$
$$\dot{q} = (A_2 + uB_2)q, \;\; y_2(t) = b^\top \Phi_2(t;u)q(0). \tag{5.4}$$

Figure 5.1 has symbolic diagrams for two types of system composition. The input–output relation of the first is $u \to y_1 + y_2$ so it is called a sum composition or parallel composition. Let $p = \mathrm{col}(x_1, \ldots, x_n)$ and $q = \mathrm{col}(x_{n+1}, \ldots, x_{2n})$. The system transition matrices are given by the bilinear system on $\mathbb{R}^{2n}$

$$\dot{x} = \begin{bmatrix} A_1 + uB_1 & 0 \\ 0 & A_2 + uB_2 \end{bmatrix} x, \quad y = \begin{bmatrix} a^\top \\ b^\top \end{bmatrix} x. \tag{5.5}$$

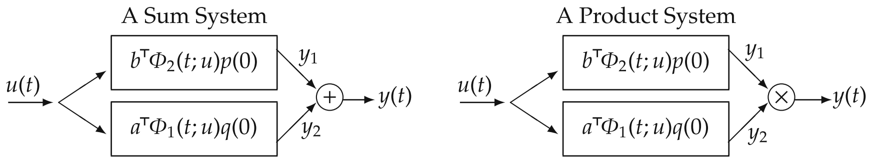


Fig. 5.1. The input–output relations $u \to y_1 + y_2$ and $u \to y_1 y_2$ have bilinear representations.

To obtain the second, a product system $u \to y_1y_2$, let $y := y_1y_2 = a^\top pq^\top b$ and $Z := pq^\top$; the result is a matrix control system in the space of rank-one matrices $\mathbf{S}_1 \subset \mathbb{R}^{n\times n}$;

$$\begin{aligned}\dot{y} =& \dot{y}_1y_2 + y_1\dot{y}_2 = a^\top \dot{p}q^\top b + a^\top p\dot{q}^\top b\\ &= a^\top\Big((A_1 + uB_1)pq^\top + pq^\top(A_2 + uB_2)^\top\Big)b;\\ \dot{Z} =& (A_1 + uB_1)Z + Z(A_2 + uB_2)^\top,\ y = a^\top Zb. \end{aligned} \tag{5.6}$$

To see that (5.6) describes the input–output relation of an ordinary bilinear system[2] on $\mathbb{R}^\nu$, rewrite it using the Kronecker product and flattened matrix notations found in Section A.3 and Exercises 1.6–1.8. For $\nu = n^2$ take as our state vector

$$x := Z^\flat = \mathrm{col}(Z_{1,1}, Z_{2,1}, \dots, Z_{n,n}) \in \mathbb{R}^\nu, \ \text{ with } x(0) = (\xi\zeta^\top)^\flat;$$

then we get a bilinear system

$$\dot{x} = (I \otimes A_1 + A_2 \otimes I)x + u(I \otimes B_1 + B_2 \otimes I)x,\ y = (ab^\top)^\flat x. \tag{5.7}$$

A special product system is $u \to y_1^2$ given the system $\dot{p} = Ap + uBp$, $y_1 = c^\top p$. Let $y = y_1^2 = c^\top pp^\top c$, $Z = pp^\top$, and $x := Z^\flat$;

$$\begin{aligned}\dot{Z} = (A + uB)Z + Z(A + uB)^\top, \text{ and using (A.9)}\\ \dot{x} = (I \otimes A + A \otimes I)x + u(I \otimes B + B \otimes I)x,\ y = (cc^\top)^\flat x.\end{aligned}$$

See Section 5.7 for more about Kronecker products.

A cascade composition of a first bilinear system $u \to v$ to a second $v \to y$, also called a series connection, yields a quadratic system, not bilinear; the relations are

$$\begin{aligned}v(t) = a^\top \Phi_1(t; u)\xi, \quad y(t) = b^\top \Phi_2(t; v)\zeta, \text{ i.e.,}\\ \dot{p} = (A_1 + uB_1)p,\ v = a^\top p,\\ \dot{q} = (A_2 + a^\top pB)q,\ y = b^\top q.\end{aligned}$$

Given two discrete-time bilinear systems $\overset{+}{p} = (A_1 + uB_1)p$, $y_1 = a^\top p$; $\overset{+}{q} = (A_2 + uB_2)q$, $y_2 = b^\top q$ on $\mathbb{R}^n$, that their parallel composition $u \to y_1 + y_2$ is a bilinear system on $\mathbb{R}^{2n}$ can be seen by a construction like (5.5). For a discrete-time product system, the mapping $u \to y_1y_2$ can be written as a product of two polynomials and is quadratic in u; the product system's transition mapping is given using the matrix $Z := pq^\top \in \mathbf{S}_1$:

[2] For a different treatment of the product bilinear system, see Remark 8.1 about shuffle products.

$$\overset{+}{Z} = (A_1 + uB_1)Z(A_2 + uB_2)^\top, \; y = (ab^\top)^\flat Z^\flat.$$

5.2 Observability

In practice, the state of a bilinear control system is not directly observable; it must be deduced from p output measurements, usually assumed to be linear: $y = Cx$, $C \in \mathbb{R}^{p\times n}$. There are reasons to try to find $x(t)$ from the histories of the inputs and the outputs, respectively,

$$\mathfrak{U}_t := \{u(\tau), \tau \le t\}, \; \mathfrak{Y}_t := \{y(\tau), \tau \le t\}.$$

One such reason is to construct a state-dependent feedback control as in Chapter 3; another is to predict, given future controls, future output values $y(T)$, $T > t$. The initial state ξ is assumed here to be unknown.

For continuous- and discrete-time bilinear control systems on $\mathbb{R}^n_*$, we use the transition matrices $X(t;u)$ defined in Chapter 1 for $u \in \mathcal{PK}$ and $P(k;u)$ defined in Chapter 4 for discrete-time inputs. Let

$$A_u := A + \sum_{i=1}^{m} u_i B_i;$$

$$\dot{x} = A_u x, \;\; u \in \mathcal{LI}, \; y(t) := CX(t;u)\xi, \; t \in \mathbb{R}_+; \tag{5.8}$$

$$\overset{+}{x} = A_u x, \;\; y(t) := CP(t;u)\xi, \; t \in \mathbb{Z}_+. \tag{5.9}$$

Assume that A, B_i, C are known and that u can be freely chosen on a time interval $[0, T]$. The following question will be addressed, beginning with its simplest formulations: *For what controls u is the mapping $\xi \mapsto \{y(t) | 0 \le t \le T\}$ one-to-one?*

5.2.1 *Constant-Input Problems*

We begin with the simplest case, where $u(t) \in \mathbb{R}^m$ is a constant $(\mu_1, \dots, \mu_m)$; then either (5.8) or (5.9) becomes a linear time-invariant dynamical system with p outputs. (To compare statements about the continuous and discrete cases in parallel the symbol $\|$ is used.)

$$\dot{x} = A_\mu x, \; y = Cx \in \mathbb{R}^p; \; \| \quad \overset{+}{x} = A_\mu x, \; y = Cx; \tag{5.10}$$

$$y(t) = Ce^{tA_\mu}\xi, \; t \in \mathbb{R}; \; \| \quad y(t) = CA_\mu^t \xi, \; t \in \mathbb{Z}_+$$

Here is a preliminary definition for both continuous and discrete-time cases: an *autonomous* dynamical system is said to be observable if an output history

$$\mathfrak{Y}^c := \{y(t),\ t \in \mathbb{R}\} \quad \| \quad \mathfrak{Y}^d := \{y(t), t \in \mathbb{Z}_+\}$$

uniquely determines ξ. For both cases, let

$$\mathbf{O}(A;\mu) := \mathrm{col}\left(C, CA_\mu, \ldots, CA_\mu^{n-1}\right) \in \mathbb{R}^{np \times p}.$$

Proposition 5.1. $\mathbf{O}(A;\mu)$ *has rank n if and only if* (1.18) *is observable.*

Proof. Assume first that rank $\mathbf{O}(A;\mu) = n$. In the continuous-time case, ξ is determined uniquely by equating corresponding coefficients of t in

$$y(0) + t\dot{y}(0) + \cdots \frac{t^{n-1}y^{(n-1)}(0)}{(n-1)!} = C\xi + tCA_\mu\xi + \cdots + \frac{t^{n-1}CA_\mu^{n-1}}{(n-1)!}\xi. \tag{5.11}$$

In the discrete-time case, just write the first n output values $\mathfrak{Y}^d_{n-1}$ as an element of $\mathbb{R}^{np}$; then

$$\begin{bmatrix} y(0) \\ \vdots \\ y(n-1) \end{bmatrix} = \begin{bmatrix} C\xi \\ \vdots \\ CA_\mu^{n-1}\xi \end{bmatrix} = \mathbf{O}(A;\mu)\xi. \tag{5.12}$$

For the converse, the uniqueness of ξ implies that the rank of $\mathbf{O}(A;\mu)$ is full. In (5.11) and (5.12), one needs at least $(n-1)/p$ terms and (by the Cayley–Hamilton Theorem) no more than n. □

Proposition 5.2. *Let* $p(\mu) = \det(\mathbf{O}(A;\mu))$. *If* $p(\mu) \neq 0$ *then equation* (5.11) [*or* (5.12)] *can be solved for* ξ, *and* $\{C, A_\mu\}$ *is called an* observable pair. △

The values of $\mu \in \mathbb{R}^m$ for which all $n \times n$ minors of $\mathbf{O}(A;\mu)$ vanish are affine varieties. Finding the radical that determines them requires methods of computational algebraic geometry [110] like those used in Section 2.5; the polynomials here are usually not homogeneous.

5.2.2 Observability Gram Matrices

The Observability Gram matrices to be defined below in (5.13) are an essential tool in understanding observability for (5.8) and (5.9). They are, respectively, an integral $\mathbf{W}^c_{t;u}$ and a sum $\mathbf{W}^d_{t;u}$ for system (5.8) with $u \in \mathcal{PK}$, and system (5.9) with any control sequence.

$$\mathbf{W}^c_{t;u} := \int_0^t Z^c_{s;u}\,ds, \qquad \| \ \mathbf{W}^d_{t;u} := \sum_{j=0}^{j} Z^d_{u;j}, \quad \text{where}$$
$$Z^c_{s;u} := X^\top(s;u)C^\top CX(s;u), \qquad \| \ Z^d_{t;u} := P(t;u)^\top C^\top CP(t;u), \quad \text{so}$$
$$\mathbf{W}^c_{t;u}\xi = \int_0^t X^\top(s;u)C^\top y(s)\,dt; \| \ \mathbf{W}^d_{t;u}\xi = \sum_{j=0}^{t} P(j;u)^\top C^\top y(j). \tag{5.13}$$

The integrand $Z^c_{s;u}$ and summand $Z^d_{t;u}$ are solutions, respectively, of

$$\dot{Z}^c = A_u^\top Z A_u,\ Z^c(0) = C^\top C, \| \ \overset{+}{Z}{}^d = A_u^\top Z^d A_u,\ Z^d(0) = C^\top C.$$

The Gram matrices are positive semidefinite and

$$\mathbf{W}^c_{t;u} \gg 0 \Longrightarrow \quad \xi = (\mathbf{W}^c_{t;u})^{-1} \int_0^t X^\top(s;u)cy(s)\,ds, \tag{5.14}$$
$$\mathbf{W}^d_{t;u} \gg 0 \Longrightarrow \quad \xi = (\mathbf{W}^d_{t;u})^{-1} \sum_{j=0}^{t} P(j;u)^\top cy(j). \tag{5.15}$$

Definition 5.1. System (5.8) [system (5.9)] is observable if there exists a piecewise constant u on any interval $[0,T]$ such that $\mathbf{W}^c_{T;u} \gg 0$, $[\mathbf{W}^d_{T;u} \gg 0]$. △

In each case Proposition 1.3, Sylvester's criterion, can be used to check a given u for positive definiteness of the Gram matrix; and given enough computing power one can search for controls that satisfy that criterion; see Example 5.1.

Returning to constant controls $u(t) \equiv \mu$

$$\mathbf{W}^c_{t;\mu} := \int_0^t \mathrm{e}^{s(A_\mu)^\top} C^\top C \mathrm{e}^{sA_\mu}\,ds; \quad \| \quad \mathbf{W}^d_{t;\mu} := \sum_{t=0}^{t} (A_\mu^\top)^t C^\top C (A_\mu)^t.$$

The assumption that for some μ the origin is stable is quite reasonable in the design of control systems. In the continuous-time case, suppose A_μ is a Hurwitz matrix; then by a standard argument (see [65]) the limit

$$\mathbf{W}^c_\mu := \lim_{t\to\infty} \mathbf{W}^c_{t;\mu} \text{ exists and } A_\mu^\top \mathbf{W}^c_\mu + \mathbf{W}^c_\mu A_\mu = -C^\top C.$$

In the discrete-time case, if the Schur stability test $|\mathrm{spec}(A_\mu)| < 1$ is satisfied then there exists

$$\mathbf{W}^d_\mu := \lim_{t\to\infty} \mathbf{W}^d_{t;\mu}; \quad A_\mu^\top \mathbf{W}^d_\mu A_\mu - W^d_\mu = -C^\top C.$$

Example 5.1. This example is unobservable for every constant control μ, but becomes observable if the control has at least two constant pieces. Let $\dot{x} = A_u x$, $x(0) = \xi$, $y = Cx$ with $C = (1, 0, 0)$ and let

$$A_u := \begin{bmatrix} 0 & 1 & u \\ -1 & 0 & 0 \\ u & 0 & 0 \end{bmatrix}; \text{ if } u(t) \equiv \mu \text{ then } \mathbf{O}(A;\mu) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \mu \\ \mu^2 - 1 & 0 & 0 \end{bmatrix}$$

has rank 2 for all μ (unobservable). Now use a $\mathcal{PK}$ control with $u(t) = \pm 1$; on each piece $A_u^3 = 0$ so $\exp(tA_u) = I + tA_u + A_u^2 t^2/2$. The system can be made observable by using only two segments:

$$u(t) = 1,\ t \in [0, 1), \quad u(t) = -1,\ t \in (1, 2].$$

The transition matrices and output are

$$e^{tA_1} = \begin{bmatrix} 1 & t & t \\ -t & 1 - t^2/2 & -t^2/2 \\ t & t^2/2 & 1 + t^2/2 \end{bmatrix},\ e^{tA_{-1}} = \begin{bmatrix} 1 & t & -t \\ -t & 1 - t^2/2 & t^2/2 \\ -t & -t^2/2 & 1 + t^2/2 \end{bmatrix};$$

$$y(0) = \xi_1,\ y(1) = \xi_1 + \xi_2 + \xi_3,\ y(2) = -\xi_1 + \xi_2 - \xi_3. \tag{5.16}$$

The observations (5.16) easily determine ξ. △

5.2.3 *Geometry of Observability*

Given an input $u \in \mathcal{U}$ on the interval $[0, T]$, call two initial states $\bar{\xi},\ \hat{\xi} \in \mathbb{R}^n$ u-indistinguishable on $[0, T]$ and write

$$\bar{\xi} \sim_u \hat{\xi} \text{ provided } CX(t;u)\bar{\xi} = CX(t;u)\hat{\xi},\ 0 \le t \le T.$$

This relation $\sim_u$ is transitive, reflexive, and symmetric (an equivalence relation) and linear in the state variables. From linearity, one can write this relation as $\bar{\xi} - \hat{\xi} \sim_u 0$, so the only reference state needed is 0. The observability Gram matrix is non-decreasing in the sense that if $T > t$ then $\mathbf{W}_{T;u} - \mathbf{W}_{t,u}$ is positive semidefinite. Therefore, rank $\mathbf{W}_{t;u}$ is non-decreasing. The set of states u-indistinguishable from the origin on $[0, T]$ is defined as

$$\mathcal{N}(T;u) := \{x|\, x \sim_u 0\} = \{x|\, \mathbf{W}_{T;u} x = 0\},$$

which is a linear subspace and can be called the u-unobservable subspace; the u-observable subspace is the quotient space

$$\mathcal{O}(T;u) = \mathbb{R}^n/\mathcal{N}(T;u);\ \mathbb{R}^n = \mathcal{N}(T;u) \oplus \mathcal{O}(T;u).$$

If $\mathcal{N}(T;u) = 0$, we say that the given system is u-observable.[3]

If two states $x,\ \xi$ are u-indistinguishable for *all* admissible $u \in \mathcal{U}$, we say that they are $\mathcal{U}$-indistinguishable and write $x \sim \xi$. The set of $\mathcal{U}$-unobservable states is evidently a linear space

$$\mathcal{N}_{\mathcal{U}} := \{x|\, \mathbf{W}_{t;u}x = 0,\ u \in \mathcal{U}\}.$$

Its quotient subspace (complement) in $\mathbb{R}^n$ is $\mathcal{O}_{\mathcal{U}} := \mathbb{R}^n/\mathcal{N}$ and is called the $\mathcal{U}$-observable subspace; the system (5.8) is called $\mathcal{U}$-observable if $\mathcal{O}_{\mathcal{U}} = \mathbb{R}^n$. The $\mathcal{U}$-unobservable subspace $\mathcal{N}_{\mathcal{U}}$ is an invariant space for the system (5.8): trajectories that begin in $\mathcal{N}_{\mathcal{U}}$ remain there for all $u \in \mathcal{U}$. If the largest invariant linear subspace of C is 0, (5.8) is $\mathcal{U}$-observable; this property is also called *observability under multiple experiment* in [108] since to test it one would need many duplicate systems, each initialized at $x(0)$ but using its own control u.

Theorem 5.1 (Grasselli–Isidori [108]). *If $\mathcal{PK} \subset \mathcal{U}$, (5.8) is $\mathcal{U}$-observable if and only if there exists a single input u for which it is u-observable.*

Proof. The proof is by construction of a universal input $\tilde{u}$ which distinguishes all states from 0; it is obtained by the concatenation of at most $n+1$ inputs u_i of duration 1: $\tilde{u} = u^0 \star \cdots \star u^n$. Here is an inductive construction. Let $u_0 = 0$, then $\mathcal{N}_0$ is the kernel of $\mathbf{W}_{u_0,1}$.

At the kth stage in the construction, the set of states indistinguishable from 0 is $\mathbf{W}_{v,k}$ where $v = u_0 \star u_1 \star \cdots u_k$. □

The test for observability that results from this analysis is that

$$\mathbf{O}(A, B) := \mathrm{col}(C, CA, CB, CA^2, CAB, CBA, CB^2, \dots) \tag{5.17}$$

has rank n. Since there are only n^2 linearly independent matrices in $\mathbb{R}^{n\times n}$, at most n^2 rows of $\mathbf{O}(A, B)$ are needed.

Theorem 5.1 was the first on the existence of universal inputs and was subsequently generalized to C^ω nonlinear systems in Sontag [246] (discrete-time) and Sussmann [265] (continuous-time).

Returning for a moment to the topic of discretization by the method of Section 1.8, (1.51), Sontag [249, Prop. 5.2.11 and Sec. 3.4] provides an observability test for linear systems that is immediately applicable to constant-input bilinear systems.

Proposition 5.3. *If $\dot{x} = (A + uB)x,\ y = Cx$ is observable for constant $u = \mu$, then observability is preserved for $\overset{+}{x} = e^{\tau(A+\mu B)}x,\ y = Cx$ provided that $\tau(\alpha - \beta)$ is not of the form $2k\pi i$ for any pair of eigenvalues α, β of $A + \mu B$ and any integer k.* △

How can one guarantee that observability cannot be destroyed by *any* input? To derive the criterion of Williamson [287][4] find necessary conditions

[3] In Grasselli and Isidori [108], u-observability is called *observability under single experiment*.

[4] The exposition in [287] is different; also see Gauthier and Kupka [103].

using polynomial inputs $u(t) = \mu_0 + \mu_1 t + \cdots + \mu_{n-1}t^{n-1}$ as in Example 1.3. We consider the case $p = 1$, $C = c^\top$. We already know from the case $u \equiv 0$ that necessarily $\boxed{\text{(i): rank}\, \mathbf{O}(A) = n.}$

To get more conditions repeatedly differentiate $y(t) = Cx(t)$ and evaluate at $t = 0$ to obtain $\mathbf{Y} := \mathrm{col}(y_0, \dot{y}_0, \ldots, y_0^{(n-1)}) \in \mathbb{R}^{n\times p}$.

First, $\dot{y}_0 = CA\xi + \mu_0 CB\xi$; if there exists μ_0 such that $\dot{y}_0 = 0$ then $\{\xi | C(A + \mu_0)\xi = 0\} \neq 0$ is unobservable; this gives $\boxed{\text{(ii): } CB = 0.}$ Continuing this way, $\ddot{y}_0 = C(A^2 + \mu_0(AB + BA + \mu_0 B^2) + \mu_1 B)\xi$ and using (a) $\ddot{y}_0 = C(A^2 + \mu_0 AB)\xi$ because the terms beginning with $c^\top B$ vanish; so necessarily $\boxed{\text{(iii): } CAB = 0.}$ The same pattern persists, so the criterion is

$$\text{rank } \mathbf{O}(A) = n \text{ and } CA^k B = 0,\ 0 \leq k \leq n-1. \tag{5.18}$$

To show their sufficiency just note that no matter what $\mu_0, \mu_1, \ldots$ are used, they do not appear in $\mathbf{Y}$, so $\mathbf{Y} = \mathbf{O}(A)\xi$ and from (i) we have $\xi = (\mathbf{O}(A))^{-1}Y$.

5.3 State Observers

Given a control u for which the continuous-time system (5.8) is u-observable with known coefficients, by computing the transition matrices $X(t; u)$ from (5.14) it is possible to estimate the initial state (or present state) from the histories $\mathfrak{U}_T$ and $\mathfrak{Y}_T$:

$$\xi = (\mathbf{W}^c_{T;u})^{-1} \int_0^T X^\top(t; u) C^\top y(t)\, dt.$$

The present state $x(T)$ can be obtained as $X(T; u)\xi$ or by more efficient means. Even though observability may fail for any constant control, it still may be possible, using some piecewise constant control u, to achieve u-observability. Even then, the Gram matrix is likely to be badly conditioned. Its condition number (however defined) can be optimized in various ways by suitable inputs; see Example 5.1.

One recursive estimator of the present state is an asymptotic state observer, generalizing a method familiar in linear control system theory. For a given bilinear or biaffine control system with state x, the asymptotic observer is a model of the plant to be observed; the model's state vector is denoted by z and it is provided with an input proportional to the output error $e := z - x$.[5] The total system, plant plus observer, is given by the following plant, observer, and error equations; K is an $n \times p$ gain matrix at our disposal.

[5] There is nothing to be gained by more general ways of introducing the output error term; see Grasselli and Isidori [109].

$$
\begin{aligned}
&\dot{x} = Ax + u(Bx + b), \ y = Cx; \\
&\dot{z} = Az + u(Bz + b) + uK(y - Cz); \ e = z - x; \\
&\dot{e} = (A + u(B - KC))e.
\end{aligned} \tag{5.19}
$$

Such an observer is designed by finding K for which the observer is convergent, meaning that $\|e(t)\| \to 0$, under some assumptions about u. Observability is more than what is needed for the error to die out; a weaker concept, detectability, will do. Roughly put, a system is detectable if what cannot be observed is asymptotically stable. At least three design problems can be posed for (5.19); the parameter identification problem for linear systems has many similarities.

1. Design an observer which will converge for all choices of u. This requires the satisfaction of the Williamson conditions (5.18), replacing B with $B - KC$. The rate of convergence depends only on spec(A).
2. Assume that u is known and fixed, in which case well-known methods of designing observers of time-variant linear systems can be employed (K should satisfy a suitable matrix Riccati equation). If also we want to choose K to get best convergence of the observer, its design becomes a nonlinear programming problem; Sen [239] shows that a random sequence of values of $u \in \mathcal{PK}$ should suffice as a universal input (see Theorem 5.1).

For observer design and stabilizing control (by output feedback) of observable inhomogeneous bilinear systems $\dot{x} = Ax + u(Bx + b)$, $y = Cx$, see Gauthier and Kupka [103], which uses the observability criteria (5.18) and the weak ad-condition; a discrete-time extension of this work is given in Lin and Byrnes [185].

Example 5.2 (Frelek and Elliott [99]). The observability Gram matrix $\mathbf{W}_{T;u}$ of a bilinear system (in continuous- or discrete-time) can be optimized in various ways by a suitable choice of input. To illustrate, return to Example 5.1 with the piecewise constant control

$$u(t) = \begin{cases} 1, & t \in [0, 1) \\ -1, & t \in (1, 2] \end{cases}.$$

Here are the integrands $Z_{1;t}$, $0 \le t \le 1$ and $Z_{-1;t}$, $1 < t \le 2$, and the values of $\mathbf{W}_{t;u}$ at $t = 1$ and $t = 2$, using

$$e^{tA_{\pm 1}} = \begin{bmatrix} 1 & t & \pm t \\ -t & 1 - \frac{t^2}{2} & \mp \frac{t^2}{2} \\ \pm t & \pm \frac{t^2}{2} & 1 + \frac{t^2}{2} \end{bmatrix}. \quad Z_{1;t} := e^{tA_1^\top} C^\top C e^{tA_1} = \begin{bmatrix} 1 & t & t \\ t & t^2 & t^2 \\ t & t^2 & t^2 \end{bmatrix},$$

$$Z_{-1;t} := e^{tA_{-1}^{\top}} e^{A_1^{\top}} C^{\top} C e^{A_1} e^{tA_{-1}} =$$

$$\begin{bmatrix} t - 2t^2 + \frac{4t^3}{3} & t - \frac{t^2}{2} - t^3 + \frac{t^4}{2} & t - \frac{3t^2}{2} + t^3 - \frac{t^4}{2} \\ t - \frac{t^2}{2} - t^3 + \frac{t^4}{2} & t + t^2 - \frac{t^3}{3} - \frac{t^4}{2} + \frac{t^5}{5} & t - \frac{t^3}{3} + \frac{t^4}{2} - \frac{t^5}{5} \\ t - \frac{3t^2}{2} + t^3 - \frac{t^4}{2} & t - \frac{t^3}{3} + \frac{t^4}{2} - \frac{t^5}{5} & t - t^2 + t^3 - \frac{t^4}{2} + \frac{t^5}{5} \end{bmatrix};$$

$$W_{1;1} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}; \ W_{u;2} = W_{1;1} + \int_1^2 Z_{-1;t}\, dt = \begin{bmatrix} \frac{16}{3} & \frac{1}{2} & \frac{-7}{2} \\ \frac{1}{2} & \frac{7}{10} & \frac{3}{10} \\ \frac{-7}{2} & \frac{3}{20} & \frac{121}{30} \end{bmatrix}$$

whose leading principal minors are positive; therefore, $W_{u;2} \gg 0$. The condition number (maximum eigenvalue)/(minimum eigenvalue) of $W_{u;2}$ can be minimized by searching for good piecewise inputs; that was suggested by Frelek [99] as a way to assure that the unknown ξ can be recovered as uniformly as possible. △

5.4 Identification by Parameter Estimation

By a black box, I will mean a control system with m inputs, p outputs, and unknown transition mappings. Here it will be assumed that $m = 1 = p$. Given a black box (either discrete- or continuous-time), the system identification problem is to find an adequate mathematical model for it. In many engineering and scientific investigations, one looks first for a linear control system that will fit the observations; but when there is sufficient evidence to dispose of that possibility a bilinear or biaffine model may be adequate.

The first possibility to be considered is that we also know n. Then for biaffine systems ($b \neq 0$) the standard assumption in the literature is that $\xi = 0$ and that the set of parameters $\{A, B, c^{\top}, b\}$ is to be estimated by experiments with well-chosen inputs. In the purely bilinear case, $b = 0$ and the parameter set is $\{A, B, c^{\top}, \xi\}$. One can postulate an arbitrary value $\xi = \delta_1$ (or $b = \delta_1$) for instance; for any other value, the realization will be the same up to a linear change of coordinates.

For continuous-time bilinear and biaffine systems, Sontag et al. [250] show the existence of a set of quadruples $\{A, B, c^{\top}, b\}$, generic in the sense of Remark 2.5, which can be distinguished by the set of variable-length pulsed inputs (with $\alpha \neq 0$)

$$u_{(\tau,\alpha)}(t) = \begin{cases} \alpha, & 0 \leq t < \tau \\ 0, & t \geq \tau \end{cases}.$$

In the world of applications, parameter estimation of biaffine systems from noisy input–output data is a statistical problem of considerable difficulty, even in the discrete-time case, as Brunner and Hess [44] point out. For several parameter identification methods for such systems using $m \geq 1$ mutually independent white noise inputs, see Favoreel et al. [88]. Gibson et al. [104] explore with numerical examples a maximum likelihood method for identifying the parameters using an EM (expectation–maximization) algorithm.

5.5 Realization

An alternative possibility to be pursued in this section is to make no assumptions about n but suppose that we know an input–output mapping accurately. The following realization problem has had much research over many years and is a literature that deserves a fuller account than I can give. For either $T = \mathbb{R}_+$ or $T = \mathbb{Z}_+$, one is given a mapping $\mathcal{H} : \mathfrak{U} \to \mathfrak{Y}$ on $[0, \infty)$, assumed to be strictly causal: $\mathfrak{Y}_T := \{y(t) | 0 \leq t \leq T\}$ is independent of $\mathfrak{U}_T := \{u(t) | T < t\}$. The goal is to find a representation of $\mathcal{H}$ as a mapping $\mathfrak{U} \to \mathfrak{Y}$ defined by system (5.8) or system (5.9). Note that the *relation* between input and output for bilinear systems is not a mapping unless an initial state $x(0) = \xi$ is given.

Definition 5.2. A homogeneous bilinear system on $\mathbb{R}^n_*$ is called span-reachable from ξ if the linear span of the set attainable from ξ is $\mathbb{R}^n$. However, a biaffine system is called span-reachable[6] if the linear span of the set attainable from the origin is $\mathbb{R}^n$. △

Definition 5.3. The realization of $\mathcal{H}$ is called canonical if it is span-reachable and observable. There may be more than one canonical realization. △

5.5.1 *Linear Discrete-Time Systems*

For a simple introduction to realization and identification, we use linear time-invariant discrete-time systems[7] on $\mathbb{R}^n$ with $m = 1 = p$, always with $x(0) = 0 = y(0)$. It is possible (by linearity, using multiple experiments) to obtain this so-called *zero-state response*; for $t > 0$, the relation between input and output becomes a mapping $y = \mathcal{H}u$. Let $\mathbf{z}$ be the unit time-delay operator, $\mathbf{z}f(t) := f(t-1)$. If

[6] Span-reachable is called *reachable* in Brockett [35], d'Allessandro et al. [217], Sussmann [262], and Rugh [228].

[7] The realization method for continuous-time time-invariant linear systems is almost the same: the Laplace transform of (1.20) on page 10 is $\hat{y}(s) = c^\top(sI - A)^{-1}\hat{u}(s)b$; its Laurent expansion $h_1 s^{-1} + h_2 s^{-2} + \cdots$ has the same Markov parameters $c^\top A^k b$ and Hankel matrix as (5.20).

$$\overset{+}{x} = Ax + ub, \; y = c^\top x, \quad \text{then}$$

$$y(t+1) = c^\top \sum_{i=0}^{t} A^i u(t-i)b, \; t \in \mathbb{Z}_+; \text{ so} \tag{5.20}$$

$$y = c^\top (I - \mathbf{z}A)^{-1} bu. \tag{5.21}$$

Call $(A, b, c^\top)$ an $\mathbb{R}^n$-triple if $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}^n$. The numbers $h_i := c^\top A^i b$ are called the Markov parameters of the system. The input–output mapping is completely described by the rational function

$$h(\mathbf{z}) := c^\top (I - \mathbf{z}A)^{-1} b; \quad y(t+1) = \sum_{i=1}^{t} h_i u(t-i).$$

The series $h(\mathbf{z}) = h_0 + h_1\mathbf{z} + h_2\mathbf{z}^2 + \cdots$ is called the *impulse response* of (5.20) because it is the output corresponding to $u(t) = \delta(t, 0)$. The given response is a rational function and (by the Cayley–Hamilton Theorem) is proper in the sense that

$$h = p/q \text{ where } \deg(p) < \deg(q) = n; \; q(\mathbf{z}) = \mathbf{z}^n + q_1\mathbf{z}^{n-1} + q_2\mathbf{z}^{n-2} + q_n.$$

If we are given a proper rational function $h(\mathbf{z}) = p(\mathbf{z})/q(\mathbf{z})$ with q of degree n then there exists an $\mathbb{R}^n$-triple for which (5.20) has impulse response h; system (5.20) is called a realization of h. For linear systems, controllability and span-reachability are synonymous. A realization that is controllable and observable on $\mathbb{R}^n$ realization is minimal (or canonical). For linear systems, a minimal realization is unique up to a nonsingular change of coordinates. There are other realizations for larger values of n; if we are given (5.20) to begin with, it may have a minimal realization of smaller dimension.

In the realization problem, we are given a sequence $\{h_i\}_1^\infty$ without further information, but there is a classical test for the rationality of $h(\mathbf{z})$ using the truncated Hankel matrices

$$H_k := \{h_{i+j-1} | \, 1 \le i, j \le k\}, \; k \in \mathbb{N}. \tag{5.22}$$

Suppose a finite number n is the least integer such that $\operatorname{rank} H_n = \operatorname{rank} H_{n+1} = n$. Then the series $h(\mathbf{z})$ represents a rational function. There exist factorizations $H_{n+1} = LR$ where $R \in \mathbb{R}^{n,n+1}$ and $L \in \mathbb{R}^{n+1,n}$ are of rank n. Given one such factorization $(LT^{-1})(TR)$ is another, corresponding to a change of coordinates. We will identify R and L with, respectively, controllability (see (1.43)) and observability matrices by seeking an integer n and a triple $A, b, c^\top$ for which

$$L = \mathbf{O}(c^\top; A), \; R = \mathbf{R}(A; b).$$

In that way, we will recover a realization of (5.21).

Suppose for example that we are given H_2 and H_3, both of rank 2. The following calculations will determine a two-dimensional realization $\{A, b, c^\top\}$.[8] Denote the rows of H_3 by $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$; compute the coefficients α_1, α_0 such that $\mathbf{h}_3 + \alpha_1\mathbf{h}_2 + \alpha_0\mathbf{h}_1 = 0$, which must exist since $\operatorname{rank} H_2 = 2$. ($A^2 + \alpha_1 A + \alpha_0 I = 0$.) Solve $LR = H_3$ for 3×2 matrix L and 2×3 matrix R, each of rank 2. Let $c^\top = l_1$, the first row of L, and $b = r_1$, the first column of R. To find A (with an algorithm we can generalize for the next section), we need a matrix M such that $M = LAR$; if

$$M := \begin{bmatrix} h_{2,1} & h_{2,2} \\ h_{3,1} & h_{3,2} \end{bmatrix}, \quad \text{then } A = (L^\top L)^{-1} L^\top M R (RR^\top)^{-1}.$$

5.5.2 Discrete Biaffine and Bilinear Systems

The input–output mapping for the biaffine system

$$\overset{+}{x} = Ax + a + \sum_{i=1}^{m} u_i(t)(B_i x + b_i), \;\; y = c^\top x, \;\; \xi = 0; \;\; t \in \mathbb{Z}_+ \tag{5.23}$$

is linear in $b_1, \dots, b_m$. Accordingly, realization methods for such systems were developed in 1972–1980 by generalizing the ideas in Section 5.5.1.[9]

A biaffine realization (5.23) of an input–output mapping may have a state space of indefinitely large dimension; it is called minimal if its state space has the least dimension of any such realization. Such a realization is minimal if and only if it is span-reachable and observable. This result and differing algorithms for minimal realizations of biaffine systems are given in Isidori [139], Tarn and Nonoyama [275, 276], Sontag [245], and Frazho [98].

For the homogeneous bilinear case treated in the examples below, the initial state ξ may not be known, but appears as a parameter in a linear way. The following examples may help to explain the method used in Tarn and Nonoyama [215, 275], omitting the underlying minimal realization theory and arrow diagrams. Note that the times and indices of the outputs and Hankel matrices start at $t = 0$ in these examples.

Example 5.3. Suppose that (B, ξ) is a controllable pair. Let

$$x(t+1) = (I + u(t)B)x, \;\; y = c^\top x, \;\; x(0) = \xi, \text{ and } h(j) := c^\top B^j \xi, \;\; j = 0, 1, \dots.$$

---

[8] The usual realization algorithms in textbooks on linear systems do not use this $M = LR$ factorization. The realization is sought in some canonical form such as $c = \delta_1$, $b = \operatorname{col}(h_{1,1}, \dots, h_{n,1})$ (the first column of H), and A the companion matrix for $t^n + \alpha_1 t^{n-1} + \cdots + \alpha_0$.

[9] The multi-output version requires a block Hankel matrix, given in the cited literature.

We can get a formula for y using the coefficients $v_j(t)$ defined by

$$\sum_{j=0}^{t} v_j(t)(-\lambda)^j = \prod_{i=0}^{t}(\lambda - u(i)) : \quad y(t) = \sum_{j=0}^{t-1} h(j)v_j(t); \text{ here}$$

$$v_0(t) = 1,\ v_1(t) = \sum_{0\le i\le t-1} u(i),\ v_2(t) = \sum_{\substack{0\le i\le t-2\\ i\le j\le t-1}} u(i)u(j),\ \ldots,\ v_{t-1}(t) = \prod_{i=0}^{t-1} u(i).$$

To compare this with (5.20), substitute ξ for b, B for A, and the monomials v_i for $u(i)$, we obtain the linear system (5.20) and use a shifted Hankel matrix $H := \{c^\top B^{i+j}\xi |\, i, j \in \mathbb{Z}_+\}$. △

Example 5.4. As usual, homogeneous and symmetric bilinear systems have the simplest theory. This example has $m = 2$ and no drift term. The underlying realization theory and arrow diagrams can be found in [139, 215, 275]. Let $\mathbf{s} := \{1, 2\}$, $\mathbf{k} := \{0, 1, \ldots, k\}$. If

$$\overset{+}{x} = (u_1 B_1 + u_2 B_2)x,\ y = c^\top x,\ x(0) = \xi \text{ then} \tag{5.24}$$

$$x(t+1) = \Big(u_1(t)B_1 + u(t)B_2\Big)\ldots\Big(u_1(0)B_1 + u_2(0)B_2\Big)\xi, \tag{5.25}$$

$$y(t+1) = \sum_{k=0}^{t} \sum_{\substack{i_j\in\mathbf{s}\\ j\in\mathbf{k}}} \Big(c^\top B_{i_k}\ldots B_{i_1}B_{i_0}\xi\Big)\Big(u_{i_1}(0)u_{i_2}(1)\ldots u_{i_k}(k)\Big) \tag{5.26}$$

$$= \sum_{k=0}^{t} \sum_{\substack{i_j\in\mathbf{s}\\ j\in\mathbf{k}}} \tilde{f}_{i_1,\ldots,i_k}\Big(u_{i_1}(0)u_{i_2}(1)\ldots u_{i_k}(k)\Big), \tag{5.27}$$

where $\tilde{f}_{i_1,\ldots,i_k} := c^\top B_{i_k}\xi \ldots B_{i_1}B_{i_0}\xi$. We now need a (new) observability matrix $\mathbf{O} := \mathbf{O}(B_1, B_2) \in \mathbb{R}^{\infty\times n}$ and a *reachability matrix* $\mathbf{R} \in \mathbb{R}^{n\times\infty}$, shorthand for $\mathbf{R}(B_1, B_2)$. We also need the first k rows $\mathbf{O}_k$ from $\mathbf{O}$ and the first k columns $\mathbf{R}_k$ of $\mathbf{R}$. It was shown earlier that if $\operatorname{rank}\mathbf{O}_t = n$ for some t then (5.24) is observable. The given system is span-reachable if and only if $\operatorname{rank}\mathbf{R}_s = n$ for some s. The construction of the two matrices $\mathbf{O}$ and $\mathbf{R}$ is recursive, by respectively stacking columns of row vectors and adjoining rows of column vectors in the following way[10] for $t \in \mathbb{Z}_+$ and $\eta(t) := 2^{t+1} - 1$.

---

[10] For systems with m inputs, the matrices $\mathbf{O}_{\eta(t)}$, $\mathbf{R}_{\eta(t)}$ on page 157 have $\eta(t) := m^{t+1} - 1$.

$$\mathbf{o}_0 := c^\intercal,\ \mathbf{o}_1 := \begin{bmatrix}\mathbf{o}_0 B_1\\ \mathbf{o}_0 B_2\end{bmatrix}, \dots, \mathbf{o}_t := \begin{bmatrix}\mathbf{o}_{t-1} B_1\\ \mathbf{o}_{t-1} B_2\end{bmatrix};$$

$\mathbf{O}_{\eta(t)} := \operatorname{col}(\mathbf{o}_0, \dots, \mathbf{o}_t)$. Thus

$$\mathbf{O}_1 = \mathbf{o}_0 = c^\intercal,\ \text{next}\ \mathbf{O}_3 := \begin{bmatrix}\mathbf{o}_0\\ \mathbf{o}_1\end{bmatrix} = \begin{bmatrix}c^\intercal\\ c^\intercal B_1\\ c^\intercal B_2\end{bmatrix},\ \text{and}\ \mathbf{O}_7 = \begin{bmatrix}\mathbf{o}_0\\ \mathbf{o}_1\\ \mathbf{o}_2\end{bmatrix}$$

$$= \operatorname{col}(c^\intercal, c^\intercal B_1, c^\intercal B_2, c^\intercal B_1^2, c^\intercal B_1 B_2, c^\intercal B_2 B_1, c^\intercal B_2^2).$$

$\mathbf{r}_0 := \xi,\ \mathbf{r}_1 := (B_1\mathbf{r}_0, B_2\mathbf{r}_0), \dots, \mathbf{r}_t := (B_1\mathbf{r}_{t-1}, B_2\mathbf{r}_{t-1})$;

$\mathbf{R}_{\eta(t)} := (\mathbf{r}_0, \dots, \mathbf{r}_t)$; so $\mathbf{R}_1 = \mathbf{r}_0 = \xi,\ \mathbf{R}_3 = \big(\mathbf{r}_0, \mathbf{r}_1\big) = \big(\xi, B_1\xi, B_2\xi\big)$,

$\mathbf{R}_7 = \big(\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_3\big) = \big(\xi, B_1\xi, B_2\xi, B_1^2\xi, B_1B_2\xi, B_2B_1\xi, B_2^2\xi\big)$. Let

$$\mathbf{v}_t := \operatorname{col}\Big(1, u_1(0), u_2(0), u_1(0)u_1(1), u_1(0)u_2(1), u_2(0)u_2(1), \dots, \prod_{k=0}^{t} u_2(k)\Big)$$

$$\text{then}\ \ y(t+1) = c^\intercal \mathbf{R}_{\eta(t)} \mathbf{v}_t. \tag{5.28}$$

In analogy to (5.22), this system's Hankel matrix is defined as the infinite matrix $\tilde{H} := \mathbf{OR}$; its finite sections are given by the sequence $\{\tilde{H}_{\eta(t)} := \mathbf{O}_{\eta(t)}\mathbf{R}_{\eta(t)},\ t \in \mathbb{N}\}$, $\tilde{H}_{\eta(t)}$ is of size $\eta(t) \times \eta(t)$. Its upper left $k \times k$ submatrices are denoted by $\tilde{H}_k$; they satisfy $\tilde{H}_k = \mathbf{O}_k\mathbf{R}_k$. Further, if for some r and s in $\mathbb{N}$ $\operatorname{rank}\mathbf{O}_r = n$ and $\operatorname{rank}\mathbf{R}_s = n$, then for $k = \max\{r, s\}$, $\operatorname{rank}\tilde{H}_k = n$, and $\operatorname{rank}\tilde{H}_{k+1} = n$.

For instance, here is the result of 49 experiments (31 unique ones).

$\tilde{H}_7 =$

$$\begin{bmatrix}
c^\intercal\xi & c^\intercal B_1\xi & c^\intercal B_2\xi & c^\intercal B_1^2\xi & c^\intercal B_1B_2\xi & c^\intercal B_2B_1\xi & c^\intercal B_2^2\xi\\
c^\intercal B_1\xi & c^\intercal B_1^2\xi & c^\intercal B_1B_2\xi & c^\intercal B_1^3\xi & c^\intercal B_1^2B_2 & c^\intercal B_1B_2B_1 & c^\intercal B_1B_2^2\\
c^\intercal B_2\xi & c^\intercal B_2B_1\xi & c^\intercal B_2^2\xi & c^\intercal B_2B_1^2\xi & c^\intercal B_2B_1B_2\xi & c^\intercal B_2^2B_1\xi & c^\intercal B_2^3\xi\\
c^\intercal B_1^2\xi & c^\intercal B_1^3\xi & c^\intercal B_1^2B_2\xi & c^\intercal B_1^3\xi & c^\intercal B_1^2B_2\xi & c^\intercal B_1^2B_2B_1\xi & c^\intercal B_1^2B_2^2\xi\\
c^\intercal B_1B_2\xi & c^\intercal B_1B_2B_1\xi & c^\intercal B_1B_2^2\xi & c^\intercal B_1B_2B_1^2\xi & c^\intercal B_1B_2B_1B_2\xi & c^\intercal B_1B_2^2B_!\xi & c^\intercal B_1B_2^3\xi\\
c^\intercal B_2B_1\xi & c^\intercal B_2B_1^2 & c^\intercal B_2B_1B_2\xi & c^\intercal B_2B_1^3\xi & c^\intercal B_2B_1^2B_2\xi & c^\intercal B_2B_1B_2B_1\xi & c^\intercal B_2B_1B_2^2\xi\\
c^\intercal B_2^2\xi & c^\intercal B_2^2B_1\xi & c^\intercal B_2^3\xi & c^\intercal B_2^2B_1^2\xi & c^\intercal B_2^2B_1B_2\xi & c^\intercal B_2^3B_!\xi & c^\intercal B_2^4\xi
\end{bmatrix}$$

Now suppose we are given a priori an input–output mapping — a sequence of outputs expressed as multilinear polynomials in the controls, where the indices $\{i_1, \dots, i_k\}$ on coefficient α correspond to the controls used:

$$y(0) = \alpha_0;$$
$$y(1) = \alpha_1 u_1(0) + \alpha_2 u_2(0);$$
$$y(2) = \alpha_{1,1}u_1(1)u_1(0) + \alpha_{1,2}u_1(1)u_2(0) + \alpha_{2,1}u_2(1)u_1(0) + \alpha_{2,2}u_2(1)u_2(0);$$

$$y(3) = \alpha_{1,1,1}u_1(0)u_1(1)u_1(2) + \alpha_{1,1,2}u_1(0)u_1(1)u_2(2) + \cdots$$
$$+ \alpha_{2,2,2}u_2(0)u_2(1)u_2(2); \cdots$$
$$y(t+1) = \sum_{k=0}^{t} \sum_{\substack{i_j \in \mathbf{s} \\ j \in \mathbf{k}}} \alpha_{i_1,\dots,i_k}\Big(u_{i_1}(0)u_{i_2}(1)\dots u_{i_k}(k)\Big). \tag{5.29}$$

These coefficients $\alpha_{i_1,\dots,i_k}$ are the entries in the numerical matrices H_T; an entry $\alpha_{i_1,\dots,i_k}$ corresponds to the entry $\tilde{f}_{i_1,\dots,i_k}$ in $\tilde{H}$. For example

$$H_1 := \begin{bmatrix}\alpha_0\end{bmatrix},\ H_2 := \begin{bmatrix}\alpha_0 & \alpha_1 \\ \alpha_1 & \alpha_2\end{bmatrix},\ H_3 := \begin{bmatrix}\alpha_0 & \alpha_1 & \alpha_2 \\ \alpha_1 & \alpha_{1,1} & \alpha_{2,1} \\ \alpha_2 & \alpha_{1,2} & \alpha_{2,2}\end{bmatrix}, \dots$$

$$H_7 := \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_{1,1} & \alpha_{2,1} & \alpha_{1,2} & \alpha_{2,2} \\ \alpha_1 & \alpha_{1,1} & \alpha_{2,1} & \alpha_{1,1,1} & \alpha_{2,1,1} & \alpha_{1,2,1} & \alpha_{2,2,1} \\ \alpha_2 & \alpha_{1,2} & \alpha_{2,2} & \alpha_{1,1,2} & \alpha_{2,1,2} & \alpha_{1,2,2} & \alpha_{2,2,2} \\ \alpha_{1,1} & \alpha_{1,1,1} & \alpha_{2,1,1} & \alpha_{1,1,1} & \alpha_{2,1,1} & \alpha_{1,2,1,1} & \alpha_{2,2,1,1} \\ \alpha_{2,1} & \alpha_{1,2,1} & \alpha_{2,2,1} & \alpha_{1,1,2,1} & \alpha_{2,1,2,1} & \alpha_{1,2,2,1} & \alpha_{2,2,2,1} \\ \alpha_{1,2} & \alpha_{1,1,2} & \alpha_{2,1,2} & \alpha_{1,1,1,2} & \alpha_{2,1,2,2} & \alpha_{1,2,1,2} & \alpha_{2,2,1,2} \\ \alpha_{2,2} & \alpha_{1,2,2} & \alpha_{2,2,2} & \alpha_{1,1,2,2} & \alpha_{2,1,2,2} & \alpha_{1,2,2,2} & \alpha_{2,2,2,2} \end{bmatrix}$$

As an exercise, one can begin with the simplest interesting case of (5.24) with $n = 2$, and generate $y(0)$, $y(1)$, $y(2)$ as multilinear polynomials in the controls; their coefficients are used to form the matrices H_0, H_1, H_2. In our example, suppose now that $\operatorname{rank} H_3 = 2$ and $\operatorname{rank} H_4 = 2$. Then any of the factorizations $H_3 = \bar{L}\bar{R}$ with $\bar{L} \in \mathbb{R}^{3\times 2}$, and $\bar{R} \in \mathbb{R}^{2\times 3}$, can be used to obtain a realization $\{\bar{c}^\top, \bar{B}_1, \bar{B}_2, \bar{\xi}\}$ on $\mathbb{R}^2$, as follows.

The 3×2 and 2×3 matrices

$$\bar{L} = \operatorname{col}(l_1, l_2, l_3),\ \bar{R} = \begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix}$$

can be found with Mathematica by

```
Sol=Solve[L.R==H, Z][[1]],
```

where the list of unknowns is `Z` $:= \{l_{1,1}, \dots l_{3,2}, r_{1,1}, \dots, r_{3,2}\}$. The list of substitution rules `Sol` has two arbitrary parameters that can be used to specify $\bar{L}, \bar{R}$. Note that the characteristic polynomials of B_1 and B_2 are preserved in any realization.

Having found $\bar{L}$ and $\bar{R}$, take $\bar{\xi} := l_1$, $\bar{c}^\top := r_1$. The desired matrices $\bar{B}_1, \bar{B}_2$ can be calculated with the help of two matrices M_1, M_2 that require data from $y(3)$ (only the three leftmost columns of H_7 are needed):

$$M_1 = \begin{bmatrix} \alpha_1 & \alpha_{1,1} & \alpha_{2,1} \\ \alpha_{1,1} & \alpha_{1,1,1} & \alpha_{2,1,1} \\ \alpha_{1,2} & \alpha_{1,1,2} & \alpha_{2,1,2} \end{bmatrix},\ M_2 = \begin{bmatrix} \alpha_2 & \alpha_{1,2} & \alpha_{2,2} \\ \alpha_{2,1} & \alpha_{1,2,1} & \alpha_{2,2,1} \\ \alpha_{2,2} & \alpha_{1,2,2} & \alpha_{2,2,2} \end{bmatrix};$$

they correspond, respectively, to $\mathbf{O}_1 B_1 \mathbf{R}_1$ and $\mathbf{O}_1 B_2 \mathbf{R}_1$. The matrices $\bar{B}_1, \bar{B}_2$ are determined uniquely from

$$\bar{B}_1 = (\bar{L}^\top \bar{L})^{-1} \bar{L}^\top M_1 R^\top (RR^\top)^{-1},\ \bar{B}_2 = (\bar{L}^\top \bar{L})^{-1} \bar{L}^\top M_2 R^\top (RR^\top)^{-1}.$$

Because the input–output mapping is multilinear in $\mathfrak{U}$, as pointed out in [275] the only needed values of the controls $u_1(t), u_2(t)$ are zero and one in examples like (5.24). △

Example 5.5. The input–output mapping for the single-input system

$$\overset{+}{x} = B_1 x + u_2 B_2 x,\ y = c^\top x,\ x(0) = \xi \tag{5.30}$$

can be obtained by setting $u_1(t) \equiv 1$ in Example 5.4; with that substitution the mapping (5.27), the Hankel matrix sections H_T, and the realization algorithm are the same. The coefficients correspond by the evaluation $u_1 = 1$, for instance the constant term in the polynomial $y(t+1)$ for the system (5.30) corresponds to the coefficient of $u_1(0) \dots u_1(t)$ in the output (5.29) for Example 5.4. △

Exercise 5.1. Find a realization on $\mathbb{R}^2$ of the two-input mapping[11]

$$\begin{aligned}
&y(0) = 1, \quad y(1) = u_1(0) + u_2(0),\\
&y(2) = u_1(0)u_1(1) - 2u_1(1)u_2(0) + u_1(0)u_2(1) + u_2(0)u_2(1),\\
&y(3) = u_1(0)u_1(1)u_1(2) + u_1(1)u_1(2)u_2(0) - 2u_1(0)u_1(2)u_2(1) +\\
&\quad 4u_1(2)u_2(0)u_2(1) + u_1(0)u_1(1)u_2(2) - 2u_1(1)u_2(0)u_2(2) +\\
&\quad u_1(0)u_2(1)u_2(2) + u_2(0)u_2(1)u_2(2).
\end{aligned}$$

△

5.5.3 *Remarks on Discrete-Time Systems**

See Fliess [92] to see how the construction sketched in Example 5.4 yields a system observable and span-reachable on $\mathbb{R}^n$ where n is the rank of the Hankel matrix, and how to classify such systems. It is pointed out in [92] that the same input–output mapping can correspond to an initialized homogeneous system or to the response of a biaffine system initialized at 0, using the homogeneous representation of a biaffine system given in Section 3.8. Power series in noncommuting variables, which are familiar in the theory of formal languages, were introduced into control theory in [92]. For their application in the continuous-time theory, see Chapter 8.

Discrete-time bilinear systems with m inputs, p outputs, and fixed ξ were investigated by Tarn and Nonoyama [215, 275] using the method of Example

[11] Hint: the characteristic polynomials of B_1, B_2 should be $\lambda^2 - 1$, $\lambda^2 + \lambda - 2$ in either order.

5.4. They obtained span-reachable and observable biaffine realizations of input–output mappings in the way used in Sections 3.8 and 5.5.2: embed the state space as a hyperplane in $\mathbb{R}^{n+1}$. Tarn and Nonoyama [215, 275, 276] uses affine tensor products in abstract realization theory and a realization algorithm.

The input–output mapping of a discrete-time biaffine system (5.23) at $\xi = 0$ is an example of the polynomial response maps of Sontag [244]; the account of realization theory in Chapter V of that book covers biaffine systems among others. Sontag [247] gave an exposition and a mathematical application to Yang-Mills theory of the input–output mapping for (4.17).

The Hankel operator and its finite sections have been used in functional analysis for over a century to relate rational functions to power series. They were applied to linear system theory by Kalman, [154, Ch. 10] and are explained in current textbooks such as Corless and Frazho [65, Ch. 7]. They have important generalizations to discrete-time systems that can be represented as machines (adjoint systems) in a category; see Arbib and Manes [8] and the works cited there.

A category is a collection of objects and mappings (arrows) between them. The objects for a *machine* are an input semigroup $\mathfrak{U}^*$, state space $\mathfrak{X}$, and output sequence $\mathfrak{Y}^*$. Dynamics is given by iterated mappings $\mathfrak{X} \to \mathfrak{X}$. (For that reason, continuous-time dynamical or control systems cannot be treated as machines.) The examples in [8] include finite-state automata, which are machines in the category defined by discrete sets and their mappings; $\mathfrak{U}$ and $\mathfrak{Y}$ are finite alphabets and $\mathfrak{X}$ is a finite set, such as Moore or Mealy machines. Other examples in [8] include linear systems and bilinear systems, using the category of linear spaces with linear mappings. Discrete-time bilinear systems "occur naturally as adjoint systems" [8] in that category. Their Hankel matrices (as in Section 5.5.2 and [276]) describe input-to-output mappings for fixed initial state ξ.

5.6 Volterra Series

One useful way of representing a causal[12] nonlinear functional $\boldsymbol{F} : \mathbb{R}_+ \times \boldsymbol{\mathcal{U}} \to \boldsymbol{\mathcal{Y}}$ is as a Volterra expansion. At each $t \in \mathbb{R}_+$

$$y(t) = K_0(t) + \sum_{i=1}^{\infty} \int_0^t \cdots \int_0^{\sigma_{i-1}} K_i(t, \sigma_1, \ldots, \sigma_i) u(\sigma_1) \cdots u(\sigma_i)\, d\sigma_i \cdots d\sigma_1. \tag{5.31}$$

The support of kernel K_k is the set $0 \le t_1 \cdots t_k \le t$ which makes this a Volterra expansion of triangular type; there are others. A kernel K_j is called separable

[12] The functional $\boldsymbol{F}(t; u)$ is causal if depends only on $\{u(s) | 0 \le s \le t\}$.

if it can be written as a finite sum of products of functions of one variable and called stationary if $K_j(t, \sigma_1, \dots, \sigma_j) = K_j(0, \sigma_1 - t, \dots, \sigma_j - t)$.

By the Weierstrass theorem, a real function on a compact interval can be approximated by polynomials and represented by a Taylor series; in precise analogy a continuous function on a compact function space $\mathcal{U}$ can, by the Stone–Weierstrass Theorem, be approximated by polynomial expansions and represented by a Volterra series.[13]

The Volterra expansions (5.31) for bilinear systems are easily derived[14] for any m; the case $m = 1$ will suffice, and $\mathcal{U} = \mathcal{PC}[0, T]$. The input–output relation $\boldsymbol{F}$ for the system

$$\dot{x} = Ax + u(t)Bx, \ x(0) = \xi, \ y = Cx$$

can be obtained from the Peano–Baker series solution (see (1.33)) for its transition matrix:

$$\begin{aligned} y(t) = C\mathrm{e}^{tA}\xi + \int_0^t u(t_1)C\mathrm{e}^{(t-t_1)A}B\mathrm{e}^{t_1 A}\xi\, dt_1 + \\ \int_0^t \int_0^{t_1} u(t_1)u(t_2)C\mathrm{e}^{(t-t_1)A}B\mathrm{e}^{(t_1-t_2)A}B\mathrm{e}^{t_2 A}\xi\, dt_2\, dt_1 + \cdots . \end{aligned} \tag{5.32}$$

This is formally a stationary Volterra series whose kernels are exponential polynomials in t. For any finite T and β such that $\sup_{0 \le t \le T} |u(t)| \le \beta$, the series (5.32) converges on $[0, T]$ (to see why, consider the Taylor series in t for $\exp(B \int_0^t u(s)ds)$. Using the relation $\exp(tA)B\exp(-tA) = \exp(\mathrm{ad}_A)(B)$, the kernels K_i for the series (5.32) can be written

$$\begin{aligned} K_0(t) &= C\mathrm{e}^{At}\xi, \\ K_1(t, \sigma_1) &= C\mathrm{e}^{tA}\mathrm{e}^{-\sigma_1 \mathrm{ad}_A}(B)\xi, \\ K_2(t, \sigma_1, \sigma_2) &= C\mathrm{e}^{tA}\mathrm{e}^{-\sigma_1 \mathrm{ad}_A}(B)\mathrm{e}^{-\sigma_2 \mathrm{ad}_A}(B)\xi, \ \cdots . \end{aligned}$$

Theorem 5.2 (Brockett [38]). *A finite Volterra series has a time-invariant bilinear realization if and only if its kernels are stationary and separable.*

Corollary 5.1. *The series* (5.32) *has a finite number of terms if and only if the associative algebra generated by* $\{\mathrm{ad}_A^k(B), 0 \le k \le \nu - 1\}$ *is nilpotent.*

The realizability of a stationary separable Volterra series as a bilinear system depends on the factorizability of certain behavior matrices; early treatments by Isidori et al. [69, 141] are useful, as well as the books of Isidori

[13] See Rugh [228] or Isidori [140, Ch. 3]. In application areas like Mathematical Biology, Volterra expansions are used to represent input–output mappings experimentally obtained by orthogonal-functional (Wiener and Schetzen) methods.

[14] The Volterra kernels of the input–output mapping for bilinear and biaffine systems were given by Bruni et al. [42, 43].

[140] and Rugh [228]. For time-invariant bilinear systems, each Volterra kernel has a proper rational multivariable Laplace transform; see Rugh [228], where it is shown that compositions of systems such as those illustrated in Section 5.1 can be studied in both continuous and discrete-time cases by using their Volterra series representations.

For discrete-time bilinear systems expansions as polynomials in the input values (discrete Volterra series) can be obtained from (4.7). For $\overset{+}{x} = Ax + uBx$ with $x(0) = \xi$ and an output $y = Cx$, for each $t \in \mathbb{Z}_+$

$$y(t) = CA^t\xi + \sum_{i=1}^{t} \left(CA^{t-i}BA^{i-1}\xi\right)u(i) + \cdots + \left(CB^t\xi\right)\prod_{i=1}^{t} u(i), \tag{5.33}$$

a multilinear polynomial in u of degree t. For a two-input example see (4.7). The sum of two such outputs is again multilinear in u. It has, of course, a representation like (5.5) as a bilinear system on $\mathbb{R}^{2n}$, but there may be realizations of lower dimension.

Frazho [98] used Fock spaces, shift operators, a transform from Volterra expansions to formal power series, and operator factorizations to obtain a realization theory for discrete-time biaffine systems on Banach spaces.

5.7 Approximation with Bilinear Systems

A motive for the study of bilinear and biaffine systems has been the possibility of approximating more general nonlinear control systems. The approximations (1.21)–(1.25), are obtained by neglecting higher-degree terms. Theorem 8.1 in Section 8.1 is the approach to approximation of input–output mappings of Sussmann [261]. This section describes some approximation methods that make use of Section 5.7.

Proposition 5.4. [15] *A bilinear system $\dot{x} = Ax + uBx$ on $\mathbb{R}^n$ with a single polynomial output $y = p(x)$ such that $p(0) = 0$ can be represented as a bilinear system of larger dimension with linear output.*

Proof. Use Proposition A.2 with the σ notation of Section A.3.4. Any polynomial p of degree d in $x \in R^n$ can be written, using subscript j for a coefficient of $x^{\otimes j}$,

$$p(x) = \sum_{i=1}^{d} c_i x^{\otimes i},\ c_i : \mathbb{R}^{n^i} \to \mathbb{R};\ \text{let } z^{(d)} := \operatorname{col}(x, x^{\otimes 2}, \ldots, x^{\otimes d}),\ \text{then}$$

[15] For Proposition 5.4 see Brockett [35, Th. 2]; see Rugh [228]. Instead of the Kronecker powers, state vectors that are pruned of repetitions of monomials can be used, with loss of notational clarity.

$$\dot{z}^{(d)} = \left[A + uB\ \sigma(A + uB, 1)\ \ldots\ \sigma(A + uB, d-1)z^{(d)}\right];\ y = \left[c_1\ \ldots\ c_d\right] z^{(d)}.$$

Since $\sigma(A + uB, k) = \sigma(A, k) + u\sigma(B, k)$, we have the desired bilinear system (although of excessively high dimension) with linear output. □

Example 5.6. A result of Krener [166, Th. 2 and Cor. 2] provides nilpotent bilinear approximations to order t^k of smooth systems $\dot{x} = f(x) + ug(x)$, $y = h(x)$ on $\mathbb{R}^n$ with compact Ω. The method is, roughly, to map its Lie algebra $\mathfrak{g} := \{f, g\}_{\mathcal{L}}$ to a free Lie algebra on two generators, and by setting all brackets of degree more than k to zero, obtain a nilpotent Lie algebra $\mathfrak{n}$ which has (by Ado's Theorem) a matrix representation of some dimension h. The corresponding nilpotent bilinear system on $\mathrm{GL}(h, \mathbb{R})$ has, from [164], a smooth mapping to trajectories on $\mathbb{R}^n$ that approximate the trajectories of the nonlinear system to order t^k. A polynomial approximation to h permits the use of Proposition 5.4 to arrive at a bilinear system of larger dimension with linear output and (for the same control) the desired property. △

Example 5.7 (Carleman approximation). See Rugh [228, Sec. 3.3] for the Carleman approximation of a real-analytic control system by a bilinear system. Using the subscript j for a matrix that maps $x^{\otimes j}$ to $\mathbb{R}^n$,

$$A_j \in \mathbb{R}^{n^j \times n}.\ \text{ Use the identity (A.6): }\ (A_j x^{\otimes j})(B_m x^{\otimes m}) = (A_j \otimes B_m)x^{\otimes j+m}.$$

A simple example of the Carleman method is the dynamical system

$$\dot{x} = F_1 x + F_2 x^{\otimes 2},\ y = x_1. \tag{5.34}$$

The desired linearization is obtained by the induced dynamics on $z^{(k)} := \mathrm{col}(x, x^{\otimes 2}, \ldots, x^{\otimes k})$ and $y = z_1$

$$\dot{z}^{(k)} = \begin{bmatrix} F_1 & F_2 & 0 & 0 & \ldots \\ 0 & \sigma(F_1, 1) & \sigma(F_2, 2) & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \sigma(F_1, k-1) \end{bmatrix} z^{(k)} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \sigma(F_2, k)x^{\otimes k+1} \end{bmatrix}$$

which depends on the exogenous term in $z_k^{(k+1)} = x^{\otimes k+1}$ that would have to come from the dynamics of $z^{(k+1)}$. To obtain a linear autonomous approximation, the kth Carleman approximation, that exogenous term is omitted. The unavoidable presence of such terms is called the *closure problem* for the Carleman method. The resulting error in $y(t)$ is of order t^{k+1} if F_1 is Hurwitz and $\|\xi\|$ is sufficiently small. If the system (5.34) has a polynomial output $y = p(x)$, an approximating linear dynamical system with linear output is easily obtained as in Proposition 5.4. To make (5.34) a controlled example, let $F_i := A_i + uB_i$ with bounded u. △

Chapter 6
Examples

Most of the literature of bilinear control has dealt with special problems and applications; they come from many areas of science and engineering and could profit from a deeper understanding of the underlying mathematics. A few of them, chosen for ease of exposition and for mathematical interest, will be reported in this chapter. More applications can be found in Bruni et al. [43], Mohler and Kolodziej [210], the collections edited by Mohler and Ruberti [204, 209, 227], and Mohler's books [206], [207, Vol. II].

Many bilinear control systems familiar in science and engineering have nonnegative state variables; their theory and applications, such as compartmental models in biology, are described in Sections 6.1 and 6.2. We again encounter switched systems in Section 6.3; trajectory tracking and optimal control are discussed in Section 6.4. Quantum mechanical systems and quantum computation involve spin systems that are modeled as bilinear systems on groups; this is particularly interesting because we meet complex bilinear systems for the first time, in Section 6.5.

6.1 Positive Bilinear Systems

The concentrations of chemical compounds, densities of living populations, and many other variables in science can take on only nonnegative values and controls occur multiplicatively, so positive bilinear systems are often appropriate. They appear as discrete-time models of the microeconomics of personal investment strategies (semiannual decisions about one's interest-bearing assets and debts); and less trivially, in those parts of macroeconomics involving taxes, tariffs, and controlled interest rates. For an introduction to bilinear systems in economics see d'Alessandro [68]. The work on economics reported in Brunner and Hess [44] is of interest because it points out pitfalls in estimating the parameters of bilinear models. The book

D.L. Elliott, *Bilinear Control Systems*, Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_6, © Springer Science+Business Media B.V. 2009

Dufrénot and Mignon [76] studies discrete-time noise-driven bilinear systems as macroeconomic models.

6.1.1 Definitions and Properties

There is a considerable theory of dynamical systems on convex cones in a linear space L of any dimension.

Definition 6.1. A closed convex set K (see page 21) in a linear space is called a convex cone if it is closed under *nonnegative* linear combinations:

$$K + K = K \text{ and } \mu K = K \ \text{ for all } \mu \geq 0;$$

its interior $\overset{\circ}{K}$ is an open convex cone;

$$\overset{\circ}{K} + \overset{\circ}{K} = \overset{\circ}{K} \text{ and } \mu\overset{\circ}{K} = \overset{\circ}{K} \ \text{ for all } \mu > 0. \qquad \triangle$$

To any convex cone K there corresponds an ordering $>$ on L: if $x - z \in K$ we write $x > z$. If $x > z$ implies $Ax > Az$ then the operator A is called order-preserving.

The convex cones most convenient for control work are positive orthants (quadrant, octant, etc.); the corresponding ordering is component-wise positivity. The notations for nonnegative vectors and positive vectors are respectively

$$x \geq 0: \ x_i \geq 0, \, i \in 1, \ldots, n; \quad \| \quad x > 0: \ x_i > 0, \, i \in 1, \ldots, n; \tag{6.1}$$

the nonnegative orthant $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n | \, x \geq 0\}$ is a convex cone. Its interior, an open orthant, is denoted by $\overset{\circ}{\mathbb{R}}{}^n_+$ (a notation used by Boothby [30] and Sachkov [231]). The spaces $\overset{\circ}{\mathbb{R}}{}^n_+$ are simply connected manifolds, diffeomorphic to $\mathbb{R}^n$ and therefore promising venues for control theory.

Definition 6.2. A control system $\dot{x} = f(x, u)$ for which $\mathbb{R}^n_+$ is a trap (trajectories starting in $\mathbb{R}^n_+$ remain there for all controls such that $u(\cdot) \in \Omega$) is called a positive control system. A positive dynamical system is the autonomous case $\dot{x} = f(x)$. $\triangle$

Working with positive bilinear control systems will require that some needed matrix properties and notations be defined. For $A = \{a_{i,j}\}$ and the index ranges $i \in 1, \ldots, n, \ j \in 1, \ldots, n$,

$$\begin{array}{lcll} \text{nonnegative} & \| & \text{positive} & \\ A \geq 0 : \ a_{i,j} \geq 0 & \| & A > 0 : \ a_{i,j} > 0 & (6.2) \\ A \trianglerighteq 0 : \ \text{if } j \neq i, \ a_{i,j} \geq 0, & \| & A \vartriangleright 0 : \ \text{if } j \neq i, \ a_{i,j} > 0. & (6.3) \end{array}$$

Matrices satisfying the two conditions in (6.2) are called nonnegative and positive, respectively. The two conditions defined in (6.3), where only the off-diagonal elements are constrained, are called essentially nonnegative and essentially positive, respectively. Essentially nonnegative matrices are called Metzler matrices. The two subsets of $\mathbb{R}^{n\times n}$ defined by nonnegativity ($A \geq 0$) and as Metzler matrices ($A \trianglerighteq 0$) are convex cones; the other two (open) cones defined in (6.2) and (6.3) are their respective interiors.

The simple criterion for positivity of a linear dynamical system, which can be found in Bellman [25], amounts to the observation that the velocities on the boundary of $\mathbb{R}^n_+$ must not point outside (which is applicable for any positive dynamical system).

Proposition 6.1. *Given* $A \in \mathbb{R}^{n\times n}$ *and any* $t > 0$,
(i) $\mathrm{e}^{At} \geq 0$ *if and only if A is Metzler and (ii)* $\mathrm{e}^{At} > 0$ *if and only if A is positive.*

Proof (Bellman). For "$\Rightarrow$": if any off-diagonal element $A_{i,j}, i \neq j$ is negative then $\exp(At)_{i,j} = A_{i,j}t + \cdots < 0$ for small t, which shows (i). If for any $t > 0$ all off-diagonal elements of e^{At} are zero, the same is true for A, yielding (ii).
For "$\Leftarrow$": to show the sufficiency of $A \trianglerighteq 0$, choose a scalar $\lambda > 0$ such that $A + \lambda I \geq 0$. From its definition as a series, $\exp((A + \lambda I)t) \geq 0$; also $\mathrm{e}^{\lambda I} \geq 0$ since it is a diagonal of real exponentials. Since the product of nonnegative matrices is nonnegative and these matrices all commute,

$$\mathrm{e}^{At} = \mathrm{e}^{(A+\lambda I)t-\lambda It} = \mathrm{e}^{(A+\lambda I)t}\mathrm{e}^{-\lambda It} \geq 0,$$

establishing (i). For (ii) replace "nonnegative" in the preceding sentence with "positive." □

Example 6.1. Diagonal symmetric n-input systems $\dot{x}_i = u_i x_i$, $i \in 1, \ldots, n$ (see Example 6.1, $D^+(n, \mathbb{R})$) are transitive on the manifold $\mathring{\mathbb{R}}^n_+$. They can be transformed by the substitution $x_i = \exp(z_i)$ to $\dot{z}_i = u_i$, which is the action of the Abelian Lie group $\mathbb{R}^n$ on itself. △

Beginning with Boothby [30], for the case $\Omega = \mathbb{R}$, there have been several papers about the system

$$\dot{x} = Ax + uBx \text{ on } \mathbb{R}^n_+, A \triangleright 0. \tag{6.4}$$

Proposition 6.2. *If* $\Omega = \mathbb{R}$ *a necessary and sufficient condition for* (6.4) *to be a positive bilinear system is that* $A \trianglerighteq 0$ *and B is diagonal; if* $\Omega = [0, 1]$ *then the condition is: the off-diagonal elements of B are nonnegative.*

Proof. In each case the Lie semigroup $\mathbf{S}$ is generated by products of exponential matrices $\{\mathrm{e}^{\tau A+\mu B} | \tau \geq 0, \mu \in \Omega\}$. Since $A + \mu B \trianglerighteq 0$ for each μ, a slight modification of the proof of Proposition 6.1 is enough. □

Definition 6.3. The pair (A, B) has property $\mathcal{P}$ if:
(i) A is essentially nonnegative;
(ii) B is nonsingular and diagonal, $B = \operatorname{diag}(\beta_1, \dots, \beta_n)$; and
(iii) $\beta_i - \beta_j \neq \beta_p - \beta_q$ unless $i = p, j = q$. △

Condition (iii) implies that the nonzero eigenvalues of ad_B are distinct. Property $\mathcal{P}$ is an open condition on (A, B); compare Section 2.11.

Proposition 6.3 (Boothby [30]). *For A, B with property $\mathcal{P}$, if $\operatorname{tr}(A) = 0$ then $\{A, B\}_{\mathcal{L}} = \mathfrak{sl}(n, \mathbb{R})$; otherwise $\{A, B\}_{\mathcal{L}} = \mathfrak{gl}(n, \mathbb{R})$. In either case the LARC is satisfied on $\overset{\circ}{\mathbb{R}}{}^n_+$.*

Proof. Suppose $\operatorname{tr}(A) = 0$. Since B is strongly regular (see Definition 3.4) and A has a nonzero component in the eigenspace of each nonzero eigenvalue of ad_B, as in Section 2.11 A and B generate $\mathfrak{sl}(n, \mathbb{R})$. □

6.1.2 *Positive Planar Systems*

Positive systems $\dot{x} = Ax + uBx$ on $\overset{\circ}{\mathbb{R}}{}^2_+$ with A, B linearly independent are (as in Section 3.3.3) hypersurface systems; for each ξ the set $\{\exp(tB)\xi \mid t \in \mathbb{R}\}$ is a hypersurface and disconnects $\overset{\circ}{\mathbb{R}}{}^2_+$.

The LARC is necessary, as always, but not sufficient for controllability on $\overset{\circ}{\mathbb{R}}{}^2_+$. Controllability of planar systems in the positive quadrant was shown in Boothby [30] for a class of pairs (A, B) (with properties including $\mathcal{P}$) that is open in the space of matrix pairs. The hypotheses and proof involve a case-by-case geometric analysis using the eigenstructure of $A + uB$ as u varies on $\mathbb{R}$. This work attracted the attention of several mathematicians, and it was completed as follows.

Theorem 6.1 (Bacciotti [13]). *Consider* (6.4) *on $\overset{\circ}{\mathbb{R}}{}^2_+$ with $\Omega = \mathbb{R}$. Assume that (A, B) has property $\mathcal{P}$ and $\beta_2 > 0$ (otherwise, replace u with $-u$). Let*

$$\Delta = (\beta_2 a_{11} - \beta_1 a_{22})^2 + 4\beta_1\beta_2 a_{12} a_{21}.$$

Then the system $\dot{x} = Ax + uBx$ is completely controllable on $\mathbb{R}^n_+$ if either
(i) $\beta_1 > 0$, or
(ii) $\beta_1 < 0$ but $\Delta > 0$ and $\beta_1 a_{22} - \beta_2 a_{11} > 0$.
In any other case, the system is not controllable.

Remark 6.1. The insightful proof given in [13] can only be sketched here. From Proposition 6.3 the LARC is satisfied on $\overset{\circ}{\mathbb{R}}{}^2_+$. Let

$$V(x) := x_2^{\beta_1} x_1^{\beta_2};\ \mathbf{b}V(x) = 0.$$

The level sets $\{x|\, V(x) = c\}$ are both orbits of $\mathbf{b}$ and hypersurfaces that disconnect $\mathring{\mathbb{R}}^2_+$. System (6.4) is controllable on $\mathring{\mathbb{R}}^2_+$ if and only if $\mathbf{a}V(x)$ changes sign somewhere on each of these level curves. This change of sign occurs in cases (i) and (ii) only. △

6.1.3 Systems on n-Orthants

It is easy to give noncontrollability conditions that use the existence of a trap, in the terminology of Section 3.3.2, for an open set of pairs (A, B).

Proposition 6.4 (Boothby [30, Th. 6.1]). *Suppose the real matrices A, B satisfy $A \trianglerighteq 0$ and $B = \mathrm{diag}(\beta_1, \dots, \beta_n)$. If there exists a vector $p \geq 0$ such that*

$$(i)\ \sum_{i=1}^{n} p_i\beta_i = 0 \ \textit{and}\ (ii)\ \sum_{i=1}^{n} p_i a_{i,i} \geq 0$$

then (6.4) *is not controllable on* $\mathring{\mathbb{R}}^n_+$.

Proof. Define a gauge function V, positive on $\mathring{\mathbb{R}}^n_+$ by

$$V(x) := \prod_{i=1}^{n} x_i^{p_i}; \qquad \text{if } x \in \mathbb{R}^n_+,$$

$$\dot V(x) := \mathbf{a}V + u\mathbf{b}V = \left(\sum_{i=1}^{n} p_i a_{i,i} + \sum_{\substack{i,j\in 1..n,\\ i\neq j}} (p_i x_1^{-1}) a_{i,j} x_j\right) V(x) > 0$$

by hypotheses (i,ii) and $x > 0$. Thus V increases on every trajectory. □

Sachkov [231] considered (6.4) and the associated symmetric system

$$\dot x = \sum_{i=1}^{m} u_i B_i x,\ \ \Omega = \mathbb{R}^m \tag{6.5}$$

on $\mathring{\mathbb{R}}^n_+$, with $m = n-1$ or $m = n-2$. The method is the one that we saw above in two dimensions: if (6.5) is controllable on nested level hypersurfaces $V(x) = \rho$ on the (simply connected) state space $\mathring{\mathbb{R}}^n_+$ and if the zero-control trajectories of (6.4) can cross all these hypersurfaces in both directions, then system (6.4) is globally controllable on $\mathbb{R}^n_+$.

6.2 Compartmental Models

Biaffine system models were constructed by R. R. Mohler, with several students and bioscientists, for mammalian immune systems, cell growth, chemotherapy, and ecological dynamics. For extensive surveys of biological applications see the citations of Mohler's work in the References, especially [206, 209, 210] and Volume II of [207]. Many of these works were based on compartmental models.

Compartmental analysis is a part of mathematical biology concerned with the movement of cells, molecules or radioisotope tracers between physiological compartments, which may be spacial (intracellular and extracellular compartments) or developmental as in Example 6.2. A suitable introduction is Eisen's book [79]. As an example of the use of optimal control for compartmental systems, some work of Ledzewicz and Schättler [175, 176] will be quoted.

Example 6.2 (Chemotherapy). Compartmental models in cell biology are positive systems which if controlled may be bilinear. Such a model is used in a recent study of the optimal control of cancer chemotherapy, Ledzewicz and Schättler [175, 176], compartmental models from the biomedical literature. Figure 6.1 represents a two-compartment model from these studies. The state variables are N_1, the number of tissue cells in phases where they grow and synthesize DNA; and N_2, the number of cells preparing for cell division or dividing. The mean transit time of cells from phase 1 to phase 2 is a_1. The number of cells is not conserved. Cells in phase 2 either divide and two daughter cells proceed to compartment 1 at rate $2a_2$ or are killed at rate $2ua_2$ by a chemotherapeutic agent, normalized as $u \in [0, 1]$.

The penalty terms to be minimized after a treatment interval $[0, T]$ are a weighted average $c^{\top} N(T)$ of the total number of cancer cells remaining, and the integrated negative effects of the treatment. Accordingly, the controlled dynamics and the cost to be minimized (see Section 6.4) are respectively

$$\dot{N} = \begin{bmatrix} -a_1 & 2a_2 \\ a_1 & -a_2 \end{bmatrix} N + u \begin{bmatrix} 0 & -2a_2 \\ 0 & 0 \end{bmatrix} N, \ J = c^{\top} N(T) + \int_0^T u(t)\,dt. \qquad \triangle$$

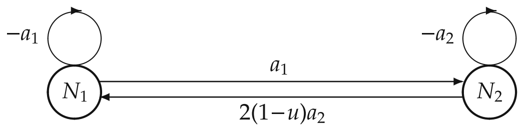


Fig. 6.1. Two-compartment model, Example 6.2.

Example 6.3. In mathematical models for biological chemistry the state variables are molar concentrations x_i of chemical species C_i, so are nonnegative. Some of the best-studied reactions involve the kinetics of the catalytic effects of enzymes, which are proteins, on other proteins. For an introduction see Banks et al. [20]. There are a few important principles used in these studies. One is that almost all such reactions occur without evolving enough heat to affect the rate of reaction appreciably. Second, the mass action law states that the rate of a reaction involving two species C_i, C_j is proportional to the rate at which their molecules encounter each other, which in turn is proportional to the product $x_i x_j$ of their concentrations measured in moles. Third, it is assumed that the molecules interact only in pairs, not trios; the resulting differential equations are inhomogeneous quadratic. These principles suffice to write down equations for $\dot{x}$ involving positive rate constants k_i that express the effects of molecular shape, energies, and energy barriers underlying the reactions.[1]

One common model describes the reaction of an enzyme E with a substrate S, forming a complex C which then decomposes to free enzyme and either a product P or the original S. Take the concentrations of S, C, and P as nonnegative state variables x_1, x_2, x_3, respectively. It is possible in a laboratory experiment to make the available concentration of E a control variable u. The resulting control system is:

$$\dot{x}_1 = -k_1 u x_1 + k_2 x_2, \quad \dot{x}_2 = k_1 u x_1 - (k_2 + k_3) x_2, \quad \dot{x}_3 = k_3 x_2. \tag{6.6}$$

It can be seen by inspection of the equations (6.6) that a conservation-of-mass law is satisfied: noting that in the usual experiment $x_2(0) = 0 = x_3(0)$, it is $x_1(t) + x_2(t) + x_3(t) = x_1(0)$. We conclude that the system evolves on a triangle in $\mathring{\mathbb{R}}^3_+$ and that its state approaches the unique asymptotically stable equilibrium at $x_1 = 0, x_2 = 0, x_3 = x_1(0)$. △

Example 6.4. A dendrite is a short (and usually branched) extension of a neuron provided with synapses at which inputs from other cells are received. It can be modeled, as in Winslow [288], by a partial differential equation (telegrapher's equation, cable equation) which can be simplified by a lumped parameter approximation. One of these, the well-known Wilfred Rall model, is described in detail in Winslow [288, Sec. 2.4]. It uses a decomposition of the dendrite into n interconnected cylindrical compartments with potentials $v := (v_1, \dots, v_n)$ with respect to the extracellular fluid compartment and with excitatory and inhibitory synaptic inputs $u := (u_1, \dots, u_n)$. The input u_i is the variation in membrane conductance between the ith compartment and potentials of 50 mv (excitatory) or −70 mv (inhibitory). The conductance between compartments i, j is $A = \{a_{i,j}\}$, which for an unbranched dendrite

[1] In enzyme kinetics the rates k_i often differ by many orders of magnitude corresponding to reaction times ranging from picoseconds to hours. Singular perturbation methods are often used to take account of the disparity of time-scales (stiffness).

is tridiagonal. The simplest bilinear control system model for the dendrite's behavior is

$$\dot{v} = Av + \sum_{1}^{n} u_j(t)N_j v + Bu(t).$$

Each $n \times n$ diagonal matrix N_j has a 1 in the (j, j) entry and 0 elsewhere, and B is constant, selecting the excitatory and inhibitory inputs for the ith compartment. The components of u can be normalized to have the values 0 and 1 only, in the simplest model. Measurements of neuronal response have been taken using Volterra series models for the neuron; for their connection with bilinear systems see Section 5.6. △

Example 6.5. This is a model of the Verhulst–Pearl model for the growth of two competing plant species, modified from Mohler [206, Ch. 5]. The state variables are populations; call them x_1, x_2. Their rates of natural growth $0 \le u_i \le 1$ for $i = 1, 2$ are under control by the experimenter, and their maximum attainable populations are $1/b_1, 1/b_2$. Here I will assume they have similar rates of inhibition, to obtain

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = u_1 \begin{bmatrix} x_1 - b_1 x_1^2 \\ -b_1 x_1 x_2 \end{bmatrix} + u_2 \begin{bmatrix} -b_2 x_1 x_2 \\ x_2 - b_2 x_2^2 \end{bmatrix}.$$

When the state is close to the origin, neglecting the quadratic terms produces a bilinear control model; in Chapter 7 a nonlinear coordinate change is produced in which the controlled dynamics becomes bilinear. △

6.3 Switching

We will say that a continuous time bilinear control system is a switching system if the control $u(t)$ selects one matrix from a given list $\mathbf{B}^m := \{B_1, \ldots, B_m\}$. (Some works on switched systems use matrices B chosen from a compact set $\mathbf{B}$ that need not be finite.) Books on this subject include Liberzon [182] and Johansson [144]. Two important types of switching will be distinguished: autonomous and controlled.

In controlled switching the controls can be chosen as planned or random $\mathcal{PK}$ functions of time. For instance, for the controlled systems discussed by Altafini [5][2]

$$\dot{x} = \sum_{i=1}^{m} u_i(t) B_i x, \;\; u_i(t) \in \{0, 1\}, \sum_{i=1}^{m} u_i(t) = 1. \tag{6.7}$$

[2] In [5] Altafini calls (6.7) a *positive scalar polysystem* because u is a positive scalar; this would conflict with our terminology.

The discrete-time version posits a rule for choosing one matrix factor at each time,

$$\overset{+}{x} = \sum_{i=1}^{m} u_i B_i x,\ u_i(t) \in \{0, 1\}, \sum_{i=1}^{m} u_i(t) = 1. \tag{6.8}$$

The literature of these systems is largely concerned with stability questions, although the very early paper Kučera [167] was concerned with optimal control of the $m = 2$ version of (6.7). For the random input case see Section 8.3.

Autonomous switching controls are piecewise linear functions of the state. The ideal autonomous switch is the sign function, partially defined as

$$\operatorname{sgn}(\phi) := \begin{cases} -1, & \phi < 0 \\ \text{undefined}, & \phi = 0 \\ 1, & \phi \geq 0 \end{cases}.$$

6.3.1 *Power Conversion*

Sometimes the nominal control is a combination of autonomous and controlled switching, with u a periodic $\mathcal{PK}$ function $\operatorname{sgn}(\sin(\omega t))$ whose on–off duty cycle is modulated by a slowly varying function ψ (pulse-width modulation):

$$u(t) = \operatorname{sgn}(\sin(\omega + \psi(t))t)\,.$$

Such controls are used for DC-to-DC electric power conversion systems ranging in size from electric power utilities to the power supply for your laptop computer. Often the switch is used to interrupt the current through an inductor, as in these examples. The control ψ regulates the output voltage.

Example 6.6. For an experimental comparison of five control methods for a bilinear model of a DC–DC boost converter see Escobar et al. [86]:

$$\dot{x}_1 = \frac{u-1}{L}x_2 + \frac{E+\epsilon}{L},\ \dot{x}_2 = \frac{1-u}{C}x_1 - \frac{1}{(R+\rho)C}x_2, \tag{6.9}$$

where the DC supply is $E + \epsilon$ volts, the load resistance is $R + \rho$ ohms, and ϵ, ρ represent parametric disturbances. The state variables are current x_1 through an inductor (L henrys) and output voltage x_2 on a C-farad capacitor; $u(t) \in \{0, 1\}$. △

Example 6.7. A simple biaffine example is a conventional ignition sparking system for an (old) automobile with a 12-volt battery, in which the primary circuit can be modeled by assigning x_1 to voltage across capacitor of value

C, x_2 to current in the primary coil of inductance L. The control is a distributor or electronic switch, either open (infinite resistance) or closed (small resistance R) in a way depending on the crankshaft angular speed ω, timing phase ψ, and a periodic function f:

$$\dot{x}_1 = \frac{-x_2}{C} - \frac{u(x_1 - 12)}{C}, \quad \dot{x}_2 = \frac{x_1}{L}, \quad u = \begin{cases} 1/R, & f(t\omega + \psi) > 0, \\ 0, & \text{otherwise.} \end{cases} \qquad \triangle$$

Other automotive applications of bilinear control systems include mechanical brakes and controlled suspension systems, and see Mohler [207].

6.3.2 *Autonomous Switching**

This subsection reflects some implementation issues for switched systems. Control engineers have often proposed feedback laws for

$$\dot{x} = Ax + uBx, \ x \in \mathbb{R}^n_*, \ \Omega = [-\mu, \mu] \tag{6.10}$$

of the form $u = \mu\,\mathrm{sgn}(\phi(x))$, where ϕ is linear or quadratic, because of their seeming simplicity. There exist (if ϕ ever changes sign) a switching hypersurface $\mathcal{S} := \{x \in \mathbb{R}^n |\, \phi(x) = 0\}$ and open regions $\mathcal{S}_+$, $\mathcal{S}_-$ where ϕ is respectively positive and negative. On each open set one has a linear dynamical system, for instance on $\mathcal{S}_+$ we have $\dot{x} = (A + \mu B)x$.

However, the definition and existence theorems for solutions of differential equations are inapplicable in the presence of state dependent discontinuities, and after trajectories reach $\mathcal{S}$ they may or may not have continuations that satisfy (6.10). One can broaden the notion of control system to permit generalized solutions. The generalized solutions commonly used to study autonomous switched systems have been Filippov solutions [90] and Krasovskii solutions [163]. They cannot be briefly described, but are as if u were rapidly chattering back and forth among its allowed values $\pm\mu$ without taking either one on an open interval. The recent article Cortés [66] surveys the solutions of discontinuous systems and their stability.

If the vector field on both sides of a hypersurface $\mathcal{S}$ points toward $\mathcal{S}$ and the resulting trajectories reach $\mathcal{S}$ in finite time, the subsequent trajectory either terminates as in the one dimensional example $\dot{x} = -\mathrm{sgn}(x)$ or can be continued in a generalized sense and is called a sliding mode. Sliding modes are intentionally used to achieve better control; see Example 6.8 for the rationale for using such variable structure systems. There are several books on this subject, such as Utkin [281]; the survey paper Utkin [280] is a good starting point for understanding sliding modes' use in control engineering. The technique in [280] that leads to useful sliding modes is simplest when the control system is given by an nth order linear differential equation. The designer chooses a switching hypersurface $\mathcal{S} := \{x \in \mathbb{R}^n | \phi(x) = 0\}$. In most applications ϕ is linear and $\mathcal{S}$ is a hyperplane as in the following example.

Example 6.8. The second-order control system $\ddot{\theta}_1 = -\operatorname{sgn}(\theta + \dot{\theta}),\ \ \theta \in S^1$, whose state space is a cylinder, arose in a work in the 1960s on vehicle roll control problems; it was at first analyzed as a planar system that is given here as a simple sliding mode example. With $x_1 := \theta$, $x_2 := \dot{\theta}$,

$$\dot{x}_1 = x_2,\ \dot{x}_2 = -\operatorname{sgn}(x_1 + x_2),\ \mathcal{S} = \{x | \, x_1 + x_2 = 0\}.$$

The portions of the trajectories that lie entirely in $\mathcal{S}_+$ or $\mathcal{S}_-$ are parabolic arcs. To show that the origin is globally asymptotically stable use the Lyapunov function $V(x) := |x_1| + x_2^2/2$, for which $\dot{V}(x) \le 0$; the origin is the only invariant set in $\{x | \, \dot{V}(x) = 0\}$, so Proposition 1.9 applies.

Each trajectory reaches $\mathcal{S}$. For large $\|\xi\|$ the trajectories switch on $\mathcal{S}$ (the trajectory has a corner but is absolutely continuous) and continue. Let $W(x) = (x_1 + x_2)^2$. $\dot{W}(x) \le 0$ in the region U where $|x_1| < 1$ and $|x_2| < 1$, so the line segment $\mathcal{S} \cap U$ is approached by all trajectories. The Filippov solution follows the line $\mathcal{S}$ toward the origin as if its dynamics were $\dot{x}_1 + x_1 = 0$. Computer simulations (Runge–Kutta method) approximate that behavior with rapid switching.

△

For bilinear control, Sira-Ramirez [242] uses Filippov solutions to study inhomogeneous bilinear models of several electrical switching circuits like the DC–DC converter (6.9). Slemrod [243] had an early version of Theorem 3.8 using $u := -\operatorname{sat}_\mu(x^\top QBx)$ as feedback and applying it to the stabilization of an Euler column in structural mechanics (for a generalization to weak stabilization of infinite-dimensional systems see Ball and Slemrod [19]). Another application of sliding mode to bilinear systems was made by Longchamp [192]. With a more mathematical approach Gauthier-Kupka [103] studied the stabilization of observable biaffine systems with control depending only on the observed output.

A Lyapunov theory for Filippov and Krasovskii solutions is given in Clarke et al. [62], which represents control systems by differential inclusions. That is, $\dot{x}$ is not specified as a function, but a closed set $K(x)$ is prescribed such that $\dot{x} \in K(x)$; for (6.10), $K(x) := Ax + \Omega Bx$. Global asymptotic stability of a differential inclusion means that the origin is stable for all possible generalized solutions and that they all approach the origin.

In practice, true sliding modes may not occur, just oscillations near $\mathcal{S}$. That is because a control transducer (such as a transistor, relay, or electrically actuated hydraulic valve) has finite limits, may be hysteretic and may have an inherent time interval between switching times. The resulting controlled dynamics is far from bilinear, but it is worthwhile to briefly describe two types of transistor models. Without leaving the realm of differential equations one can specify feedback laws that take into account the transistor's limited current by using saturation functions $u(x) = \operatorname{sat}_\delta(\phi(x))$, with parameter $\delta > 0$:

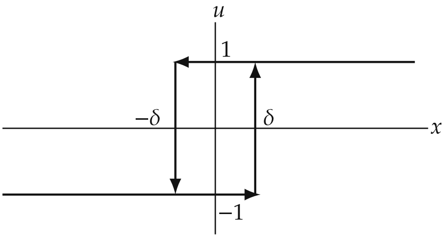


Fig. 6.2. Hysteretic switching.

$$\operatorname{sat}_\delta(\phi) := \begin{cases} \operatorname{sgn}(\phi/\delta) & |\phi| > \delta \\ \phi/\delta, & |\phi| < \delta \end{cases}. \tag{6.11}$$

A one dimensional example of switching with hysteresis is illustrated in Fig. 6.2, where $\dot{x} = -u$. The arrows indicate the sign of $\dot{x}$ and the jumps in u prescribed when x leaves the sets $(-\infty, \delta)$ and $(-\delta, \infty)$ as shown.

The definition of hybrid systems is still evolving, but in an old version the state space is $\{1, \ldots, m\} \times R^n$; the hybrid state is $(j(t), x(t))$. There is given a list of matrices $\{B_1, \ldots, B_m\}$ and a covering of $\mathbb{R}^n$ by open sets $U_1, \ldots, U_m$ with $\dot{x} = B_j x$ when $x \in U_j$. The discrete component $j(t)$ changes when x leaves U_j (as in Fig. 6.2); to make the switching rule well defined one needs a small delay between successive changes.

6.3.3 Stability Under Arbitrary Switching

The systems to be considered here are (6.7) and its discrete-time analog (6.8). Controllability problems for these systems cannot be distinguished from the general bilinear control problems. The stability problems are first the possibility of finding a stabilizing control, which is covered by Chapters 3 and 4; second, guaranteeing stability or asymptotic stability of the system no matter what the sequence of switches may be, as well as the construction of Lyapunov functions common to all the dynamical systems $\dot{x} = B_j x$. There are probabilistic approaches that will be described in Section 8.3. A recent survey of stability criteria for switched linear systems is Shorten et al. [241], which deals with many aspects of the problem, including its computational complexity.

Definition 6.4. By uniform exponential stability (UES[3]) of (6.7) we mean that uniformly for arbitrary switching functions u there exists $\delta > 0$ such that all trajectories satisfy $\|x(t)\| \leq \|x(0)\| \exp(-t\delta)$. △

[3] UES is also called global uniform exponential stability, GUES.

A necessary condition for UES of (6.7) is that each of the B_i is Hurwitz. It is notable that this condition is far from sufficient.

Example 6.9. The two matrices A, B given below are both Hurwitz if $\alpha > 0$; but if they are used in alternation for unit times with $0 < \alpha < 1/2$ the resulting transition matrices are not Schur–Cohn; let

$$A = \begin{bmatrix} -\alpha & 1 \\ -4 & -\alpha \end{bmatrix}, \; B = \begin{bmatrix} -\alpha & -4 \\ -1 & -\alpha \end{bmatrix}; \tag{6.12}$$

$$\text{if } \alpha = \frac{1}{2}, \; \operatorname{spec}(e^A e^B) \approx \{-1.03448, -0.130824\}. \qquad \triangle$$

Lemma 6.1. *If matrix $A \in \mathbb{F}^{n\times n}$ is nonsingular, $a, b \in \mathbb{F}^n$, $\alpha \in \mathbb{F}$ then*

$$\det\left(\begin{bmatrix} A & a \\ b^\top & \alpha \end{bmatrix}\right) = \det(A)(\alpha - b^\top A^{-1} a). \tag{6.13}$$

Proposition 6.5 (Liberzon et al. [184]). *In switched system* (6.7) *suppose that each matrix in $\mathbf{B}^m := \{B_1, \dots, B_m\}$ is Hurwitz and that the Lie algebra $\mathfrak{g} := \mathbf{B}^m_{\mathcal{L}}$ is solvable; then system* (6.7) *is UES and there is a common quadratic positive definite Lyapunov function V that decreases on every possible trajectory.*

Proof. By the theorem of Lie given in Varadarajan [282, Th. 3.7.3] (see Section 3.2.2), by the hypothesis on $\mathfrak{g}$ there exists a matrix $S \in \mathbb{F}^{n\times n}$ such that the matrices $C_i := -S^{-1} B_i S$ are upper triangular for $i \in 1, \dots, m$. Changing our usual habit, let the entries of C_i be denoted by $c_i^{k,l}$. Since all of the matrices in $\mathbf{B}^m$ are Hurwitz there is a $\delta > 0$ such that $\Re(c_i^{k,k}) > \delta$ for $k \in 1, \dots, m$. We need to find by induction on n a diagonal positive definite matrix $Q = \operatorname{diag}(q_1, \dots, q_n)$ such that $C_i^* Q + Q C_i \gg 0$. Remark 1.2 can be used to simplify the argument in [184] and it suffices to treat the case $n = 3$, from which the pattern of the general proof can be seen. Let D_i, E_i, F_i denote the leading principal submatrices of orders $1, 2, 3$ (respectively) of $C_i^* Q + Q C_i$ and use Lemma 6.1 to find the leading principal minors:

$$D_i := 2q_1 \Re(c_i^{1,1}), \quad E_i := \begin{bmatrix} 2q_1 \Re(c_i^{1,1}) & q_1 c_i^{1,2} \\ q_1 \bar{c}_i^{1,2} & 2q_2 \Re(c_i^{2,2}) \end{bmatrix},$$

$$F_i := \begin{bmatrix} 2q_1 \Re(c_i^{1,1}) & q_1 c_i^{1,2} & q_1 c_i^{1,3} \\ q_1 \bar{c}_i^{1,2} & 2q_2 \Re(c_i^{2,2}) & q_2 c_i^{2,3} \\ q_1 \bar{c}_i^{1,3} & q_2 \bar{c}_i^{2,3} & 2q_3 \Re(c_i^{3,3}) \end{bmatrix};$$

$$\det(D_i) = D_i = 2q_1 \Re(c_i^{1,1}) > 0;$$

$$\det(E_i) = D_i \left(2q_2 \Re(c_i^{2,2}) - q_1^2 |c_i^{1,2}|^2\right)$$

$$\det(F_i) = \det(E_i)\left(2q_3 \Re(c_i^{3,3}) - (q_1^2 |c_i^{1,3}|^2 + q_2^2 |c_i^{2,3}|^2)\right).$$

Choose $q_1 := 1$; then all $D_i > 0$. Choose

$$q_2 > \max_i \left\{ \frac{|c_i^{1,2}|^2}{2\Re(c_i^{2,2})} \right\}, \text{ then } \det(E_i) > 0,\ i \in 1, \dots, m;$$
$$q_3 > \max_i \left\{ \frac{|c_i^{1,3}|^2 + q_2^2 |c_i^{2,3}|^2}{2\Re(c_i^{3,3})} \right\}, \text{ then } \det(F_i) > 0,\ i \in 1, \dots, m.$$

Sylvester's criterion is now satisfied, and there exists a single Q such that $C_i^* Q + QC_i \gg 0$ for all $i \in 1, \dots, m$. Returning to real coordinates, let $V(x) := x^\top SQS^{-1}x$; it is a Lyapunov function that decreases along all trajectories for arbitrary switching choices, establishing UES of the system (6.7). (The theorem in [184] is obtained for a compact set of matrices $\mathbf{B}$.) □

For other methods of constructing polynomial Lyapunov functions for UES systems as approximations to smooth Lyapunov functions see Mason et al. [202]; the degrees of such polynomials may be very large.

Definition 6.5. The system (6.7) is called uniformly stable (US) if there is a quadratic positive definite Lyapunov function V that is non-increasing on each trajectory with arbitrary switching. △

Proposition 6.6 (Agrachev and Liberzon [1]). *For the switched system* (6.7) *suppose that each matrix in* $\mathbf{B}^m$ *is Hurwitz and that* $\mathbf{B}^m$ *generates a Lie algebra* $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$, *where* $\mathfrak{g}_1$ *is a solvable ideal in* $\mathfrak{g}$ *and* $\mathfrak{g}_2$ *is compact (corresponds to a compact Lie group). Then* (6.7) *is US.* △

Remark 6.2. In Proposition 6.6, [1] obtained the uniform stability property, but not the UES property, by replacing the hypothesis of solvability in Proposition 6.5 with a weaker condition on the Lie algebra; it is equivalent to nonnegative definiteness of the Cartan–Killing form, $\chi(X, Y) := \operatorname{tr}(\operatorname{ad}_X \operatorname{ad}_Y)$, on $(\mathfrak{g}, \mathfrak{g})$. For a treatment of stability problems for switched linear systems using Lyapunov exponents see Wirth [289]. △

For discrete-time systems (6.8) the definition of UES is essentially the same as for continuous time systems. If each B_i is Schur–Cohn and the Lie algebra $\mathfrak{g} := \{B_1, \dots, B_m\}_{\mathcal{L}}$ is solvable then we know that the matrices of system (6.8) can be simultaneously triangularized. According to Agrachev and Liberzon [1] one can show UES for such a switched system using a method analogous to Proposition 6.5.

There exist pairs of matrices that are both Schur–Cohn, cannot be simultaneously triangularized over $\mathbb{C}$, and the order $u(t)$ of whose products can be chosen in (6.8) so that $\|x(t)\| \to \infty$.[4] The matrices $\exp(A)$ and $\exp(B)$ of Example 6.9 is one such pair; here is another.

[4] Finding these examples is a problem of finding the worst case, which might be called pessimization, if one were so inclined.

Example 6.10. This is from Blondel–Megretski [26] Problem 10.2.

$$A_0 = \alpha\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix},\ A_1 = \alpha\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix},\ \operatorname{spec}(A_0) = \{\alpha, \alpha\} = \operatorname{spec}(A_1);$$

$$A_1 \text{ and } A_2 \text{ are Schur–Cohn if } |\alpha| < 1. \quad A_0A_1 = \alpha^2\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix};$$

$$\operatorname{spec}(A_0A_1) = \alpha^2\frac{3 \pm \sqrt{5}}{2};\ A_0A_1 \text{ is unstable if } \alpha^2 > \frac{2}{3 + \sqrt{5}} \approx 0.618. \qquad \triangle$$

The survey in Margaliot [198] of stability of switched systems emphasizes a method that involves searching for the "most unstable" trajectories (a version of the old and famous *absolute stability* problem of control theory) by the methods of optimal control theory.

6.4 Path Construction and Optimization

Assume that $\dot{x} = Ax + \sum_{i=1}^{m} u_iB_ix$ is controllable on $\mathbb{R}^n$. There is then a straightforward problem of trajectory construction:

Given ξ and ζ, find $T > 0$ and $u(t)$ on $[0, T]$ such that that $X(T; u)\xi = \zeta$.
This is awkward, because there are too many choices of u that will provide such a trajectory. Explicit formulas are available for Abelian and nilpotent systems. Consider Example 1.8 as $\dot{z} = iz + uz$ on $\mathbb{C}$. The time $T = \arg(\zeta) - \arg(\xi) + 2k\pi$, $k \in \mathbb{Z}_+$. Use any u such that

$$\zeta = e^{r(T)}e^{iT}\xi \text{ where } r(T) := \int_0^T u(t)\,dt.$$

The trajectory construction problem can (sometimes) be solved when $\dot{x} = Ax + ubc^\top x$. This rank-one control problem reduces, on the complement of the set $L = \{x|\, c^\top x = 0\}$, to the linear problem $\dot{x} = Ax + bv$ where $v = uc^\top x$ (also see its discrete-time analog (4.10)). As in Exercise 1.12, if A, b satisfy the Kalman rank criterion (1.43), given states ξ, ζ use a control v of the form $v(t) = b^\top \exp(-tA^\top)d$. Then a unique d exists such that the control $v(t) = b^\top \exp(-tA^\top)d$ provides a trajectory from ξ to ζ. If this trajectory nears L, $u(t) = v(t)/c^\top x$ grows without bound. That difficulty is not insuperable; if $(c^\top, A)$ satisfies the Kalman rank condition (5.2) better results have been obtained by Hajek [114]. Brockett [39] had already shown that if $c_i^\top\xi \neq 0$, $i \in 1, \dots, m$, for small T the attainable set $\mathcal{A}^T(\xi)$ is convex for systems $\dot{x} = Ax + \sum_{i=1}^{m} u_ib_ic_i^\top x$ with several rank-one controls.

6.4.1 Optimal Control

Another way to make the set of trajectories smaller is by providing a cost and constraints; then one can find necessary conditions for a trajectory to have minimal cost while satisfying the constraints; these conditions may make the search for a trajectory from ξ to ζ finite dimensional.

This line of thought, the optimal control approach, has engineering applications (often with more general goals than reaching a point target ζ) and is a subspecies of the calculus of variations; it became an important area of mathematical research after Pontryagin's Maximum Principle [220] was announced. R. R. Mohler began and promoted the application of optimal control methods to bilinear systems, beginning with his study of nuclear power plants in the 1960s. His work is reported in his books [205, 206] and [207, Vol. 2], which cite case studies by Mohler and many others in physiology, ecology, and engineering. Optimal control of bilinear models of pest population control was surveyed by Lee [179]. For a classic exposition of Pontryagin's Maximum Principal see Lee and Markus [177, 178]; for a survey including recent improvements see Sussmann [268]. For a history of calculus of variations and optimal control see Sussmann and Willems [272].

For linear control systems ($\dot{x} = Ax + bu$) it is known that the set of states attainable from ξ is the same for locally integrable controls with values in a finite set, say $\{-1, 1\}$, as in its convex hull, $[-1, 1]$. This is an example of the so-called bang-bang principle. It holds true for bilinear systems in which the matrices A, B commute, and some others. For such systems the computation of the optimal $u(t)$ reduces to the search for optimal switching times for a sequence of ± 1 values.

In engineering reports on the computation of time-optimal controls, as well as in Jan Kučera's pioneering studies of controllability of bilinear systems [167–169], it was assumed that one could simplify the optimization of time-to-target to this search for optimal switching times (open-loop) or finding hypersurfaces $h(x) = c$ on which a state feedback control $u = h(x)$ would switch values. However, Sussmann [256] showed that the bang-bang principle does not apply to bilinear systems in general; a value between the extremes may be required.[5]

Example 6.11. Optimal control of the angular attitude of a rigid body (a vehicle in space or under water) is an important and mathematically interesting problem; the configuration space is the rotation group SO(3). Baillieul [17] used a quadratic control cost for this problem; using Pontryagin's Maximum Principle on Lie groups it was shown that control torques around two axes suffice for controllability. $\triangle$

Example 6.12. Here is a simple optimal control problem for $\dot{x} = Ax + uBx$, $u \in \mathcal{LI}[0, T]$ with $u(t) \in \Omega$, a compact set. To assure that the accessible set is

[5] See Sussmann [257] for a counterexample to a bang-bang principle proposed in [167].

open assume the ad-condition (Definition 3.6). Given an initial state ξ and a target ζ, suppose there exists at least one control $\hat{u}$ such that $X(t;\hat{u})\xi = \zeta$. The goal is to minimize, subject to the constraints, the cost

$$J_u := \int_0^T u(t)^2\,dt. \text{ For } \epsilon > 0 \text{ the set } \{u|\, X(t;u)\xi = \zeta,\ J_u \leq J_{\hat{u}} + \epsilon\}$$

is weakly compact, so it contains an optimal control u^o that achieves the minimum of J_u. The Maximum Principle [220] provides a necessary condition on the minimizing trajectory that employs a Hamiltonian $H(x,p;u)$ where $p \in \mathbb{R}^n$ is called a costate vector.[6]

$$H(x,p;u) := \frac{u^2}{2} + p^\top(A + uB)x.$$

From the Maximum Principal, necessarily

$$H^o(x,p) := \min_u H(x,p;u) \text{ is constant;}$$

$$\text{using } u^o = -p^\top Bx,\ H^o(x,p) = p^\top(A - \frac{1}{2}Bxp^\top B)x;$$

$$\text{to find } u^o(t) \text{ and } x(t) \text{ solve}$$

$$\dot{x} = Ax - (p^\top Bx)Bx,\ \dot{p} = -A^\top p + (p^\top Bx)B^\top p,\ x(0) = \xi,\ x(T) = \zeta$$

as a two-point boundary problem to find the optimal trajectory and control. △

Example 6.13. Medical dosage problems can be seen as optimal control problems for bilinear systems. An interesting example of optimal feedback control synthesis is the pair of papers Ledzewicz and Schättler [175, 176] mentioned in Example 6.2. They are based on two- and three-dimensional compartmental models of cancer chemotherapy. The cost J is the remaining number of cancer cells at the end of the dosage period $[0,T]$. △

Exercise 6.1 (Lie and Markus [177], p. 257). Given $\dot{x} = (u_1J + u_2I)x$ with $u \in \mathcal{LI}$ constrained by $|u_1| \leq 1$, $|u_2| \leq 1$, let $\xi = \mathrm{col}(1,0)$; consider the attainable set for $0 \leq t \leq \pi$ and the target state $\zeta := -\xi$. Verify the following facts. The attainable set $\mathcal{A}_\pi(\xi)$ is neither convex nor simply connected, ζ is not on its boundary, and the control $u_1(t) = 1, u_2(t) = 0$ such that $X(\pi,u)\xi = \zeta$ satisfies the maximal principal but the corresponding trajectory (time-optimal) does not lead to the boundary of $\mathcal{A}_\pi(\xi)$.[7] △

---

[6] In the classical calculus of variations p is a Lagrange parameter and the Maximum Principal is analogous to Euler's necessary condition.

[7] In polar coordinates the system in Exercise 6.1 is $\dot{\rho} = u_2\rho$, $\dot{\theta} = u_1$, and the attainable set from ξ is $\{x|\, e^{-T} \leq \rho \leq e^T, -T \leq \theta \leq T\}$.

6.4.2 Tracking

There is a long history of path planning and state tracking[8] problems. One is to obtain the uniform approximation of a desired curve on a manifold, $\gamma : \mathbb{R} \to \mathcal{M}^n$ by a trajectory of a symmetric bilinear system on $\mathcal{M}^n$.[9] In a simple version the system is

$$\dot{X} = \sum_1^m u_i(t) B_i X, \; u \in \mathcal{PK}, \; X(0) = I, \tag{6.14}$$

and its corresponding Lie group **G** acts transitively on one of its coset spaces $\mathcal{M} := \mathbf{G}/\mathbf{H}$ where **H** is a closed subgroup of **G** (see Sections B.5.7 and 2.7). For recent work on path planning see Sussmann [267], Sussmann and Liu [270, 271], Liu [185], and Struemper [254].

Applications to vehicle control may require the construction of constrained trajectories in SO(3), its coset space $S^2 = \mathrm{SO}(3)/\mathrm{SO}(2)$ or the semidirect product $\mathbb{R}^3 \ltimes \mathrm{SO}(3)$ (see Definition B.16) that describes rigid body motions. The following medical example amused me.

Example 6.14 (Head angle control). For a common inner ear problem of elderly people (stray calcium carbonate particles, called otoliths, in the semicircular canals) that results in vertigo, there are diagnostic and curative medical procedures[10] that prescribe rotations of the human head, using the fact that the otoliths can be moved to harmless locations by the force of gravity. The patient's head is given a prescribed sequence of four rotations, each followed by a pause of about 30 s. This treatment can be interpreted as a trajectory in $\mathcal{M} := \mathbb{R}^3 \ltimes \mathrm{SO}(3)$ to be approximated, one or more times as needed, by actions of the patient or by the therapist. Mathematical details, including the numerical rotation matrices and time histories, can be found in Rajguru et al. [222]. △

Example 6.15. Here is an almost trivial state tracking problem. Suppose the m-input symmetric system (2.1) on $\mathbb{R}^n_*$ of Chapter 2 is rewritten in the form

$$\dot{x} = M(x)u \quad u \in \mathcal{PK}, \; u(\cdot) \in \Omega; \; -\Omega = \Omega, \tag{6.15}$$
$$M(x) := \begin{bmatrix} B_1 x \cdots & B_n x \cdots & B_m x \end{bmatrix}$$

is weakly transitive. Also suppose that we are given a differentiable curve $\gamma : [0, T] \to \mathbb{R}^n_*$ such that $\operatorname{rank} M(\gamma(t)) = n$, $0 \le t \le T$. The $n \times n$ matrix

[8] The related output tracking problem is, given $x(0) = \xi$, to obtain a feedback control $u = \phi(x)$ such that as $t \to \infty$ $|y(t) - r(t)| \to 0$; usually $r(t)$ is assumed to be an exponential polynomial.

[9] In some works path planning refers to a less ambitious goal: approximate the image of γ in $\mathbb{R}^n$, forgetting the velocity.

[10] For the Hallpike maneuver (diagnosis) and modified Epley maneuver (treatment) see `http://www.neurology.org/cgi/content/full/63/1/150`.

$W(x) := M(x)M^{\top}(x)$ is nonsingular on γ. If $X(t;u)\xi = \gamma(t)$, $0 \le t \le T$ then $\xi = \gamma(0)$ and

$$\dot\gamma(t) = M(\gamma(t))u(t).$$

$$\text{Define } v \text{ by } u(t) = M^{\top}(\gamma(t))v(t), \text{ then}$$
$$v = (M(\gamma(t)M^{\top}(\gamma(t)))^{-1}\dot\gamma(t); \text{ solving for } u,$$
$$\dot x = M(x)M^{\top}(\gamma(t))(M(\gamma(t)M^{\top}(\gamma(t)))^{-1}\dot\gamma(t) \tag{6.16}$$

which has the desired output. △

Example 6.16. The work of Sussmann and Liu [185, 270] clarified previous work on the approximation of Lie brackets of vector fields. In this chapter their idea can be illustrated by an example they give in [270]. Assume that

$$\dot x = u_1B_1x + u_2B_2x$$

is controllable on an open subset $U \subset \mathbb{R}^3_*$. Let $B_3 := B_2B_1 - B_1B_2$, $M(x) := \begin{bmatrix}B_1x\ B_2x\ B_3x\end{bmatrix}$ and assume that $\operatorname{rank} M(x) = 3$ on U.

Given points $\xi, \zeta \in U$ choose a smooth curve $\gamma : [0,1] \to U$ with $\gamma(0) = \xi$ and $\gamma(1) = \zeta$. By our Example 6.15 there exists an input w such that

$$\dot\gamma(t) = \Big(w_1(t)B_1 + w_2(t)B_2 + w_1(t)B_3\Big)\gamma(t).$$

So γ is a solution to the initial value problem

$$\dot x = \Big(w_1(t)B_1 + w_2(t)B_2 + w_1(t)B_3\Big)x, x(0) = \xi.$$

Defining a sequence of oscillatory controls, the kth being

$$u_1^k(t) := w_1(t) + k^{\frac12}w_3(t)\sin(kt), u_2^k(t) := w_2(t) - 2k^{\frac12}\cos(kt),$$

$$\text{let } \dot x = \Big(u_1^k(t)B_1 + u_2^k(t)B_2\Big)x,\ x(0) = \xi,$$

whose solutions $x(t;u^k)$ converge uniformly to $\gamma(t)$ as $k \to \infty$. △

Exercise 6.2. To try the preceding example let

$$B_1x := \begin{bmatrix}x_1 - x_2\\ x_1 + x_2\\ x_3\end{bmatrix},\ B_2x := \begin{bmatrix}x_3\\ 0\\ -x_1\end{bmatrix},\ B_3x := \begin{bmatrix}0\\ x_3\\ -x_2\end{bmatrix};$$
$$\det M(x) = x_3(x_1^2 + x_2^2 + x_3^2). \text{ Choose } U = \{x \mid x_3 > 0\}.$$

Choose a parametrized plane curve in U and implement (6.16) in Mathematica with `NDSolve` (see the examples); and plot the results for large k with `ParametricPlot3D`. △

6.5 Quantum Systems*

The controllability and optimal control problems for quantum mechanical systems have led to the work of interest to mathematicians versed in Lie algebras and Lie groups.

Work on quantum systems in the 1960–1980 era was mostly about quantum-limited optical communication links and quantum measurements. The controllability problem was discussed by Huang et al. [133] and a good survey of it for physicists was given by Clark [61]. The recent collaborations of D'Alessandro, Albertini, and Dahleh [2, 3, 70, 72] and Khaneja et al. [157] are an introduction to applications in nuclear magnetic resonance, quantum computation, and experimental quantum physics. There are now books (see [71]) on quantum control.

A single input controlled quantum mechanical system is described by the Schrödinger equation

$$\sqrt{-1}\hbar\dot{\boldsymbol{\psi}}(t) = (\boldsymbol{H}_0 + u\boldsymbol{H}_1)\boldsymbol{\psi}(t), \tag{6.17}$$

where $\hbar = h/(2\pi)$ is the reduced Planck's constant. It is customary to use the atomic units of physics in which $\hbar = 1$. In general the state $\boldsymbol{\psi}$ (which is a function of generalized coordinates q, p and time) is an element of a complex Hilbert space $\mathcal{H}$. The system's Hamiltonian operator is the sum of a term $\boldsymbol{H}_0$ describing the unperturbed system and an external control Hamiltonian $u\boldsymbol{H}_1$ (due, for example, to electromagnetic fields produced by the carefully constructed laser pulses discussed by Clark [61]). The operators $\boldsymbol{H}_0, \boldsymbol{H}_1$ on H are Hermitian and unbounded. They can be represented as infinite dimensional matrices, whose theory is much different from those of finite dimension. An example: for matrix Lie algebras the trace of a Lie bracket is zero, but two operators familiar to undergraduate physics students are position q and momentum $p = \frac{\partial}{\partial q}$, which in atomic units satisfy $[p, q] = I$, whose trace is infinite.

For the applications to spin systems to be discussed here, one is interested in the transitions among energy levels corresponding to eigenstates of $\boldsymbol{H}_0$; these have well established finite dimensional low-energy approximations: $\boldsymbol{\psi} \in \mathcal{H} \doteq C^n$ and the Hamiltonians have zero trace.

The observations in quantum systems are described by $|\boldsymbol{\psi}_i|^2 = \boldsymbol{\psi}_i^* \boldsymbol{\psi}_i$, $i \in 1, \dots, n$, interpreted as probabilities: $\sum_1^n |\boldsymbol{\psi}_i|^2 = 1$.[11] That is, $\boldsymbol{\psi}(t)$ evolves on the unit sphere in $\mathbb{C}^n$. Let the matrices A, B be the section in $\mathbb{C}^{n\times n}$ of $-\sqrt{-1}\boldsymbol{H}_0$ and $\sqrt{-1}\boldsymbol{H}_1$, respectively; these matrices are skew Hermitian ($A^* + A = 0$, $B^* + B = 0$) with trace zero. The transition matrix $X(t; u)$ for this bilinear system satisfies (see Section 2.10.1)

---

[11] In quantum mechanics literature the state $\boldsymbol{\psi}$ is written with Dirac's notation as a "ket" $|\psi\rangle$; a costate vector $\langle\psi|$ (a "bra") can be written $\boldsymbol{\psi}^*$; and the inner product as Dirac's *bracket* $\langle\psi_1|\psi_2\rangle$ instead of our $\boldsymbol{\psi}_1^*\boldsymbol{\psi}_2$.

$$\dot{X} = AX + uBX,\ X(0) = I;\ X^*X = I_n,\ \det(X) = 1;\ \text{so } X \in \mathrm{SU}(n). \tag{6.18}$$

Since SU(n) is compact and any two generic matrices A, B in its Lie algebra $\mathfrak{su}(n)$ will satisfy the LARC, it would seem that controllability would be almost inevitable; however, some energy transitions are not allowed by the rules of quantum mechanics. The construction of the graphs of the possible transitions (see [4] and [236]) employs the structure theory of $\mathfrak{su}(n)$ and the subgroups of SU(n).

In D'Alessandro and Dahleh [72] the physical system has two energy levels (representing the effects of coupling to spins $\pm 1/2$) as in nuclear magnetic resonance experiments; the control terms $u_1 H_1, u_2 H_2$ are two orthogonal components of an applied electromagnetic field. The operator equation (6.18) evolves on SU(2); see Example 2.27. One of the optimal control problems studied in [72] is to use two orthogonal fields to reach one of the two energy levels at time T while minimizing $\int_0^T u_1^2(t) + u_2^2(t)\, dt$ (to maintain the validity of the low-field approximation). Pontryagin's Maximum Principle on Lie groups (for which see Baillieul [17]) is used to show in the case considered that the optimal controls are the radio-frequency sinusoids that in practice are used in changing most of a population of particles from spin *down* to spin *up*; there is also a single-input problem in [72] whose optimal signals are Jacobi elliptic functions.

Second canonical coordinates have been used to obtain solutions to quantum physics problems as described in Section 2.8, especially if $\mathfrak{g}$ is solvable; see Wei and Norman [284, 285] and Altafini [6]. The groups SU(n) are compact, so Proposition 3.9 applies.

For other quantum control studies see Albertini and D'Alessandro [2, 3], D'Alessandro [70, 71], D'Alessandro and Dahleh [72], and Clark [61]. In studying chemical reactions controlled by laser action, Turinici and Rabitz [279] shows how a finite-dimensional representation of the dynamics like (6.18) can make useful predictions.

Chapter 7
Linearization

Often the phrase *linearization of a dynamical system* $\dot{x} = f(x)$ merely means the replacement of f by an approximating linear vector field. However, in this chapter it has a different meaning that began with the following question, important in the theory of dynamical systems, that was asked by Henri Poincaré [219]: *Given analytic dynamics $\dot{x} = f(x)$ on $\mathbb{F}^n$ with $f(0) = 0$ and $F := f_*(0) \neq 0$, when can one find a neighborhood of the origin* $\mathbf{U}$ *and a mapping $x = \phi(z)$ on* $\mathbf{U}$ *such that $\dot{z} = Fz$?*

Others studied this question for C^k systems, $1 \leq k \leq \infty$, which require tools other than the power series method that will be given here. It is natural to extend Poincaré's problem to these two related questions.

Question 7.1. Given a C^ω nonlinear control system

$$\dot{x} = \sum_{i=1}^{m} u_i b_i(x) \text{ with } \frac{\partial b_i(0)}{\partial x} = B_i,$$

when is there a diffeomorphism $x = \phi(z)$ such that $\dot{z} = \sum_{i=1}^{m} u_i B_i z$?

Question 7.2. Given a Lie algebra $\mathfrak{g}$ and its real-analytic vector field representation $\tilde{\mathfrak{g}}$ on $\mathbb{R}^n$, when are there C^ω coordinates in which the vector fields in $\tilde{\mathfrak{g}}$ are linear?

These and similar linearization questions have been answered in various ways by Guillemin and Sternberg [112], Hermann [124], Kušnirenko [171] Sedwick [237, 238], and Livingston [187, 188]. The exposition in Sections 7.1 and 7.2 is based on [188]. Single vector fields and two-input symmetric control systems on $\mathbb{R}^n$ are discussed here; but what is said can be stated with minor changes and proved for m-input real-analytic systems with drift on real-analytic manifolds.

We will use the vector field and Lie bracket notations of Section B.3.5: $f(x) := \mathrm{col}(f_1(x), \ldots, f_n(x))$, $g(x) := \mathrm{col}(g_1(x), \ldots, g_n(x))$, and $[f, g] := g_*(x)f(x) - f_*(x)g(x)$.

D.L. Elliott, *Bilinear Control Systems,* Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_7,

7.1 Equivalent Dynamical Systems

Two dynamical systems defined on neighborhoods of 0 in $\mathbb{R}^n$

$$\dot{x} = f(x), \quad \text{on } \mathbf{U} \subset \mathbb{R}^n \text{ and } \dot{z} = \hat{f}(z), \quad \text{on } \mathbf{V} \subset \mathbb{R}^n$$

are called C^ω equivalent if there exists a C^ω diffeomorphism $x = \phi(z)$ from $\mathbf{V}$ to $\mathbf{U}$ such that their integral curves are related by the composition law $\boldsymbol{p}_t(x) = \phi \circ \hat{\boldsymbol{p}}_t(z)$. The two dynamical systems are equivalent, in a sense to be made precise, on the neighborhoods $\mathbf{V}$ and $\mathbf{U}$.

Locally the equivalence is expressed by the fact that the vector fields f, $\hat{f}$ are ϕ-related (see Section B.7)

$$\phi_*(z)\hat{f}(z) = f \circ \phi(z) \text{ where } \phi_*(x) := \frac{\partial \phi(x)}{\partial x}. \tag{7.1}$$

The equation (7.1) is called an intertwining equation for $\mathbf{f}$ and $\mathbf{g}$. Suppose $\hat{g}$ also satisfies $\phi_*(z)\hat{g}(z) = g \circ \phi(z)$; then $\{\hat{f}, \hat{g}\}_{\mathcal{L}}$ is ϕ-related to $\{f, g\}_{\mathcal{L}}$, hence isomorphic to it.

Here are two famous theorems dealing with C^ω-equivalence. Let $\mathbf{g}$ be a C^∞ vector field on $\mathbb{R}^n$, and $\xi \in \mathbb{R}^n$. The first one is the focus of an enjoyable undergraduate textbook, Arnol'd [10].

Theorem 7.1 (Rectification). *If the real-analytic vector field $g(x)$ satisfies $g(\xi) = c \neq 0$ then there exists a neighborhood $\mathbf{U}_\xi$ and a diffeomorphism $x = \phi(z)$ on $\mathbf{U}_\xi$ such that in a neighborhood of $\zeta = \phi^{-1}(\xi)$ the image under ϕ of g is constant: $\dot{z} = a$ where $a = \phi_*(\xi)c$.*

Theorem 7.2 (Poincaré). [1] *If $g(0) = 0$ and the conditions*

$$\begin{aligned}
&(\mathrm{P}_0) \quad g_*(0) = G = P^{-1}DP \textit{ where } D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),\\
&(\mathrm{P}_1) \quad \lambda_i \notin \left\{ \sum_{j=1}^n k_i\lambda_j \,|\, k_j \in \mathbb{Z}_+, \sum_{j=1}^n k_i > 1 \right\},\\
&(\mathrm{P}_2) \quad 0 \notin K := \left\{ \sum_{i=1}^n c_i\lambda_i \,|\, c_i \in \mathbb{R}_+, \ \sum_{i=1}^n c_i = 1 \right\},
\end{aligned}$$

are satisfied, then on a neighborhood $\mathbf{U}$ of the origin the image under ϕ of $g(x)$ is a linear vector field: there exists a $\mathbb{C}^\omega$ mapping $\phi : \mathbb{R}^n \to \mathbf{U}$ such that

$$\phi_*(z)^{-1} g \circ \phi(z) = Gz. \tag{7.2}$$

[1] See [219, Th. III, p. cii]. In condition (P_2) the set K is the closed convex hull of spec(G). For the Poincaré linearization problem for holomorphic mappings ($\overset{+}{z} = f(z)$, $f(0) = 0$, $|f'(0)| \neq 1$) see the recent survey in Broer [40].

The proof is first to show that there is a formal power series[2] for ϕ, using (P_0) and (P_1); a proof of this from [188] is given in Theorem 7.5. Convergence of the series is shown by the Cauchy method of majorants, specifically by the use of an inequality that follows from (P_1) and (P_2): there exists $\kappa > 0$ such that

$$\left|\frac{\sum_j k_j\lambda_j - \lambda_1}{\sum_j k_j}\right| > \kappa.$$

Poincaré's theorem has inspired many subsequent papers on linearization; see the citations in Chen [54], an important paper whose Corollary 2 shows that the diagonalization condition (P_0) is superfluous.[3]

Condition (P_0) will not be needed subsequently. Condition (P_1) is called the non-resonance condition. Following Livingston [188], a real-analytic vector field satisfying (P_1) and (P_2) is called a Poincaré vector field. Our first example hints at what can go wrong if the eigenvalues are pure imaginaries.

Example 7.1. [4] It can be seen immediately that the normalized pendulum equation and its classical "linearization"

$$\dot{x} = \begin{bmatrix} x_2 \\ -\sin(x_1) \end{bmatrix}, \quad \dot{z} = \begin{bmatrix} z_2 \\ -z_1 \end{bmatrix}$$

are **not** related by a diffeomorphism, because the linear oscillator is isochronous (the periods of all the orbits are 2π). The pendulum orbits that start with zero velocity have initial conditions $(\alpha, 0)$, energy $(1 - \cos(\alpha)) < 2$, and period $\tau = 4K(\alpha)$ that approach infinity as $\alpha \to \pi$, since $K(\alpha)$ is the complete elliptic integral:

$$K(\alpha) = \int_0^1 \left(\left(1 - t^2\right)\left(1 - \left(\sin^2\left(\frac{\alpha}{2}\right)t^2\right)\right)^{-1/2} dt. \qquad \triangle$$

Example 7.2. If $n = 1$, $f_* := df/dx$ is continuous, and $f_*(0) = c \neq 0$ then $\dot{x} = f(x)$ is linearizable in a neighborhood of $\{0\}$. The proof is simple: the intertwining equation (7.1) becomes $y_*(x)f(x) = cy(x)$, which is linear and first order.[5] For instance, if $f(x) = \sin(x)$ let $z = 2\tan(x/2)$, $|x| < \pi/4$; then $\dot{z} = z$. $\triangle$

---

[2] From Hartman [117, IX, 13.1]: corresponding to any formal power series $\sum_{j=0}^{\infty} p^j(x)$ with real coefficients there exists a function ϕ of class C^∞ having this series as its formal Taylor development.

[3] The theorems of Chen [54] are stated for C^∞ vector fields; his proofs become easier assuming $\mathbb{C}^\omega$.

[4] From Sedwick's thesis [237].

[5] Usually `DSolve[y'[x] f(x)= cy[x], y, x]` will give a solution that can be adjusted to be analytic at $x = 0$.

7.1.1 Control System Equivalence

The above notion of equivalence of dynamical systems generalizes to nonlinear control systems; this is most easily understood if the systems have piecewise constant inputs (polysystems: Section B.3.4) because their trajectories are concatenations of trajectory arcs of vector fields.

Formally, two C^ω control systems on $\mathbb{R}^n$ with identical inputs

$$\dot{x} = u_1 f(x) + u_2 g(x),\ x(0) = x_0 \text{ with trajectory } x(t),\ t \in [0, T], \tag{7.3}$$

$$\dot{z} = u_1 \hat{f}(z) + u_2 \hat{g}(z),\ z(0) = z_0 \text{ with trajectory } z(t),\ t \in [0, T], \tag{7.4}$$

are said to be C^ω equivalent if there exists a C^ω diffeomorphism $\phi : \mathbb{R}^n \to \mathbb{R}^n$ that takes trajectories into trajectories by $x(t) = \phi(z(t))$.

Note that for each of (7.3), (7.4) one can generate as in Section B.1.6 a list of vector field Lie brackets in one-to-one correspondence to $\mathcal{PH}(m)$, the Philip Hall basis for the free Lie algebra on m generators. Let $H : \mathcal{PH}(m) \to \mathfrak{g}$, $\hat{H} : \mathcal{PH}(m) \to \hat{\mathfrak{g}}$ be the Lie algebra homomorphisms from the free Lie algebra to the Lie algebras generated by (7.3) and (7.4), respectively.

Theorem 7.3 (Krener [164]). *A necessary and sufficient condition that there exist neighborhoods $U(x_0)$, $V(z_0)$ in $\mathbb{R}^n$ and a C^ω diffeomorphism $\phi : U(x_0) \to V(z_0)$ taking trajectories of* (7.3) *with $x(0) = x_0$ to trajectories of* (7.4) *with $z(0) = z_0$ is that there exist a constant nonsingular matrix $L = \phi_*(x_0)$ that establishes the following* local isomorphism *between $\mathfrak{g}$ and $\hat{\mathfrak{g}}$:*

$$\textit{for every } p \in \mathcal{PH}(m),\ \textit{if } h := H(p),\ \hat{h} := \hat{H}(p) \textit{ then } Lh(x_0) = \hat{h}(z_0). \tag{7.5}$$

The mapping ϕ of Theorem 7.3 is constructed as a correspondence of trajectories of the two systems with identical controls; the trajectories fill up (respective) submanifolds whose dimension is the Lie rank of the Lie algebra $\mathfrak{h} := \{f, g\}_{\mathcal{L}}$. In Krener's theorem $\mathfrak{h}$ can be infinite-dimensional over $\mathbb{R}$ but for (7.5) to be useful the dimension must be finite, as in the example in [164] of systems equivalent to linear control systems $\dot{x} = Ax + ub$. Ado's Theorem and Theorem 7.3 imply Krener [166, Th. 1] which implies that a system (7.3) whose accessible set has open interior and for which $\mathfrak{h}$ is finite dimensional is *locally* a representation of a matrix bilinear system whose Lie algebra is isomorphic to $\mathfrak{h}$.[6] For further results from [166] see Example 5.6.

Even for finite-dimensional Lie algebras the construction of equivalent systems in Theorem 7.3 may not succeed in a full neighborhood of an equilibrium point x_e because the partial mappings it provides are not unique and it may not be possible to patch them together to get a single linearizing diffeomorphism around x_e. Examples in $\mathbb{R}^2$ found by Livingston [188] show that when $\mathfrak{h}$ is not transitive a linearizing mapping may, although piecewise real-analytic, have singularities.

---

[6] For this isomorphism see page 36.

Example 7.3. Diffeomorphisms ϕ for which both ϕ and ϕ^{-1} are polynomial are called birational or Cremona transformations; under composition of mappings they constitute the Cremona group $\mathbf{C}(n, \mathbb{F})$[7] acting on $\mathbb{F}^n$. It has two known subgroups: the affine transformations $\mathbf{A}(n, \mathbb{F})$ like $x \mapsto Ax + b$ and the triangular (Jonquière) subgroup $\mathbf{J}(n, \mathbb{F})$.

The mappings $\phi \in \mathbf{C}(n, \mathbb{F})$ provide easy ways to put a disguise on a real or complex linear dynamical system. Start with $\dot{x} = Ax$ and let $x = \phi(z)$; then $\dot{z} = \phi_*^{-1}(z)A\phi(z)$ is a nonlinear polynomial system. In the case that $\phi \in \mathbf{J}(n, \mathbb{F})$

$$x_1 = z_1, \quad x_2 = z_2 + p_2(z_1), \quad \cdots, \quad x_n = z_n + p_n(z_1, \quad \ldots, \quad z_{n-1}),$$

where $p_1, p_2, \ldots, p_n$ are arbitrary polynomials. The unique inverse $z = \phi^{-1}(x)$ is given by the n equations

$$z_1 = x_1,\ z_2 = x_2 - p(x_1), \ldots,\ z_n = x_n - p_n\Big(x_1, x_2 - p_2(x_1), \ldots\Big). \qquad \triangle$$

Any bilinear system

$$\dot{z} = u_1 Az + u_2 Bz \tag{7.6}$$

can be transformed by a nonlinear C^ω diffeomorphism $\psi : \mathbb{R}^n \to \mathbb{R}^n, z = \psi(x)$, with $\psi(0) = 0$, to a nonlinear system $\dot{x} = u_1 a(x) + u_2 b(x)$ where

$$a(x) := \psi_*^{-1}(x)A\psi(x), \quad b(x) := \psi_*^{-1}(x)B\psi(x). \tag{7.7}$$

Reversing this process, given a C^ω nonlinear control system on $\mathbb{R}^n$

$$\dot{x} = u_1 a(x) + u_2 b(x) \text{ with } a(0) = 0 = b(0), \tag{7.8}$$

we need hypotheses on the vector fields a and b that will guarantee the existence of a mapping ϕ, real-analytic on a neighborhood of the origin, such that (7.8) is (as in (7.7)) C^ω equivalent to the bilinear system (7.6).

One begins such an investigation by finding necessary conditions. Evaluating the Jacobian matrices tells us[8] that we may take

$$\phi_*(0) = I, \quad A = a_*(0) \text{ and } B = b_*(0).$$

Referring to Section B.3.5 and especially Definition B.7 in Appendix B, the relation (7.7) can be written as intertwining equations like

$$A\psi(x) = \psi_*(x)a(x),\ B\psi(x) = \psi_*(x)b(x); \tag{B.7$'$}$$

corresponding nodes in the matrix bracket tree $\mathcal{T}(A, B)$ and the vector field bracket tree $\mathcal{T}(\mathbf{a}, \mathbf{b})$ (analogously defined) are ϕ-related in the same way:

---

[7] The Cremona group is the subject of the Jacobian Conjecture (see [161]): if $\phi : \mathbb{C}^n \to \mathbb{C}^n$ is a polynomial mapping such that $\det(\phi_*)$ is a nonzero constant then ϕ is an isomorphism and ϕ^{-1} is also a polynomial.

[8] (7.6) is also the classical linearization (1.24) of the nonlinear system $\dot{x} = u_1 a(x) + u_2 b(x)$.

since $[A, B]\psi(x) = \psi_*(x)[a, b](x)$ and so on for higher brackets, $\mathfrak{g} := \{A, B\}_{\mathcal{L}}$ and $\tilde{\mathfrak{g}}\{a, b\}_{\mathcal{L}}$ are *isomorphic* as Lie algebras and as generated Lie algebras. So necessary conditions on (7.8) are

(H_1) the dimension ℓ of $\tilde{\mathfrak{g}}$ as is no more than n^2, and

(H_2) $f_*(0) \neq 0$ for each vector field f in a basis of $\tilde{\mathfrak{g}}$.

7.2 Linearization: Semisimplicity and Transitivity

This section gives sufficient conditions for the existence of a linearizing diffeomorphism for single vector field at an equilibrium and for a family of C^ω vector fields with a common equilibrium.[9] Since the nonlinear vector fields in view are C^ω, they can be represented by convergent series of multivariable polynomials. This approach requires that we find out more about the adjoint action of a linear vector field on vector fields whose components are polynomial in x.

7.2.1 *Adjoint Actions on Polynomial Vector Fields*

The purpose of this section is to state and prove Theorem 7.4, a classical theorem which is used in finding normal forms of analytic vector fields. For each integer $k \geq 1$ let $\mathfrak{W}^k$ denote the set of vector fields $v(x)$ on $\mathbb{C}^n$ whose n components are homogeneous polynomials of degree k in *complex* variables $x_1, \dots, x_n$. Let $\mathfrak{W}^k$ be the formal complexification of the space $\mathfrak{V}^k$ of real vector fields $v(x)$ on $\mathbb{R}^n$ with homogeneous polynomial coefficients of degree k; denote its dimension over $\mathbb{C}$ by d_k, which is also the dimension of $\mathfrak{V}^k$ over $\mathbb{R}$. It is convenient for algebraic reasons to proceed using the complex field; then at the end go to the space of real polynomial vector fields $\mathfrak{V}^k$ that we really want.

The distinct monomials of degree k in n variables have cardinality

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)!k!}, \quad \text{so } d_k = n\frac{(n+k-1)!}{(n-1)!k!}.$$

Let $\alpha := (k_1, \dots, k_n)$ be a multi-index with k_i nonnegative integers (Section B.1) of degree $|\alpha| := \sum_{j=1}^{n} k_j$. A natural basis over $\mathbb{C}$ for $\mathfrak{W}^k$ is given by the d_k column vectors

[9] Most of this section is from Livingston and Elliott [188]. The definition of ad_A there differs in sign from the ad_a used here.

$$w_j^\alpha(x) := x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n} \delta_j \text{ for } j \in 1,\dots,n, \text{ with } |\alpha| = k \text{ and as usual}$$
$$\delta_1 := \mathrm{col}(1,0,\dots,0), \quad \dots, \quad \delta_n := \mathrm{col}(0,0,\dots,1).$$

We can use a lexicographic order on the $\{w_j^\alpha\}$ to obtain an ordered basis E_k for $\mathfrak{W}^k$.

Example 7.4. For $\mathfrak{W}^2$ on $\mathbb{C}^2$ the elements of E_2 are, in lexicographic order,

$$w_1^{2,0} = \begin{bmatrix} x_1^2 \\ 0 \end{bmatrix},\ w_1^{1,1} = \begin{bmatrix} x_1 x_2 \\ 0 \end{bmatrix},\ w_1^{0,2} = \begin{bmatrix} x_2^2 \\ 0 \end{bmatrix},$$
$$w_2^{2,0} = \begin{bmatrix} 0 \\ x_1^2 \end{bmatrix},\ w_2^{1,1} = \begin{bmatrix} 0 \\ x_1 x_2 \end{bmatrix},\ w_2^{0,2} = \begin{bmatrix} 0 \\ x_2^2 \end{bmatrix}. \qquad \triangle$$

The Jacobian matrix of an element $a(x) \in \mathfrak{W}^k$ is $a_*(x) := \partial a(x)/\partial x$. An element of $\mathfrak{W}^1$ is a vector field Ax, $A \in \mathbb{C}^{n\times n}$. Given vector fields $a(x) \in \mathfrak{W}^j$ and $b(x)$ in $\mathfrak{W}^k$, their Lie bracket is given (compare (B.6)) by

$$[a(x), b(x)] := a_*(x)b(x) - b_*(x)a(x) \in \mathfrak{W}^{k+j-1}.$$

Bracketing by a first-order vector field $a(x) := Ax \in \mathfrak{W}^1$ gives rise to its adjoint action, the linear transformation denoted by

$$\mathrm{ad}_a : \mathfrak{W}^k \to \mathfrak{W}^k; \quad \text{for all } v(x) \in \mathfrak{W}^k$$
$$\mathrm{ad}_a(v)(x) := [Ax, v(x)] = Av(x) - v_*(x)Ax \in \mathfrak{W}^k. \tag{7.9}$$

Lemma 7.1. *If a diagonalizable matrix $C \in \mathbb{C}^{n\times n}$ has spectrum $\{\mu_1,\dots,\mu_n\}$ then for all $k \in \mathbb{N}$, the operator ad_d defined by $\mathrm{ad}_d(v)(x) := Cv(x) - v_*(x)Cx$ is diagonalizable on $\mathfrak{W}^k$ (using E_k) with eigenvalues*

$$\mu_r^\alpha := \mu_r - \sum_{j=1}^n k_j \mu_j,\ r \in 1,\dots,n,\ \textit{for } \alpha = (k_1,\dots,k_n),\ |\alpha| = k \tag{7.10}$$

and corresponding eigenvectors w_r^α.

Proof. Let $D := TCT^{-1}$ be a diagonal matrix; then $D\delta_r = \mu_r \delta_r$, $r \in, 1,\dots,n$, so

$$\mathrm{ad}_d(w_r^\alpha)(x) = [Dx, w_r^\alpha(x)]$$
$$= \mu_r x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n} \delta_r - \left[k_1 x_1^{k_1-1} \cdots x_n^{k_n} \delta_r \ \cdots \ k_n x_1^{k_1} \cdots x_n^{k_n-1} \delta_r\right] \begin{bmatrix} \mu_1 x_1 \\ \cdots \\ \mu_n x_n \end{bmatrix}$$
$$= \left(\mu_r - \sum_{j=1}^n k_j \mu_j\right) w_r^\alpha(x) = \mu_r^\alpha\, w_r^\alpha(x).$$

For each $w \in \mathfrak{W}^k$ we can define, again in $\mathfrak{W}^k$, $\hat{w}(x) := T^{-1}\hat{w}(Tx)$. Then $\mathrm{ad}_c(w)(x) = T^{-1}\,\mathrm{ad}_d(\hat{w})(Tx)$. So $\mathrm{ad}_d(\hat{w}) = \mu\hat{w}$ if and only if $\mathrm{ad}_c(w) = \mu w$. □

Because the field $\mathbb{C}$ is algebraically closed, any $A \in \mathbb{C}^{n\times n}$ is unitarily similar to a matrix in upper triangular form, so we may as well write this Jordan decomposition as $A = D+N$ where D is diagonal, N is nilpotent, and $[D,N] = 0$. Then the eigenvalues of A are the eigenvalues of D.

The set of complex matrices $\{A|\,a_{i,j} = 0, i < j\}$ is the upper triangular Lie algebra $\mathfrak{t}(n,\mathbb{C})$; if the condition is changed to $a_{i,j} = 0, i \leq j$, we get the strictly upper triangular Lie algebra $\mathfrak{n}(n,\mathbb{C})$. Using (7.9) we see using the basis E_k that the map

$$\psi_k : \mathfrak{t}(n,\mathbb{C}) \to \mathfrak{g}_k \subset \mathfrak{gl}(\mathfrak{W}^k), \tag{7.11}$$

$$\psi_k(A)(v) = \mathrm{ad}_a(v) \;\text{ for all } v \in \mathfrak{W}^k \tag{7.12}$$

provides a matrix representation $\mathfrak{g}_k$ of the *solvable* Lie algebra $\mathfrak{t}(n,\mathbb{C})$ on the δ_k-dimensional linear space $\mathfrak{W}^k$.

Theorem 7.4. *Suppose the eigenvalues of A are $\{\mu_1,\dots,\mu_n\}$. Then the eigenvalues of* ad_a *on* $\mathfrak{V}^k$ *are*

$$\{\mu_r^\alpha \mid |\alpha| = k,\ r \in 1,\dots,n\}$$

as in (7.10), *and under the non-resonance condition* (P_1) *none of them is zero.*

Proof. In the representation $\mathfrak{g}_k := \psi_k(\mathfrak{t}(n,\mathbb{C}))$ on $\mathfrak{W}^k$, by Lie's theorem on solvable Lie algebras (Theorem B.3) there exists some basis for $\mathfrak{W}^k$ in which the matrix $\psi_k(A)$ is upper triangular:
(a) $\mathrm{ad}_d = \psi_k(\mathrm{ad}_D)$ is a $\delta_k \times \delta_k$ complex matrix, diagonal by 7.1;
(b) $\psi_k(A) = \psi_k(D) + \psi_k(N)$ where $[\psi_k(N), \psi_k(D)] = 0$ by the Jacobi identity;
(c) $N \in [\mathfrak{t}(n,\mathbb{C}), \mathfrak{t}(n,\mathbb{C})]$, so $\mathrm{ad}_N = \psi_k(N) \in [\mathfrak{g}_k, \mathfrak{g}_k]$ must be strictly upper triangular, hence nilpotent; therefore
(d) the Jordan decomposition of ad_A on $\mathfrak{W}^k$ is $\mathrm{ad}_A = \mathrm{ad}_D + \mathrm{ad}_N$;
and the eigenvalues of $\psi(D)$ are those of $\psi(A)$. That is, $\mathrm{spec}(\mathrm{ad}_a) = \mathrm{spec}(\mathrm{ad}_d)$, so the eigenvalues of ad_a acting on $\mathfrak{W}^d$ are given by (7.10).

When A is real, $\mathrm{ad}_A : \mathfrak{V}^d \to \mathfrak{V}^d$ is linear and has the same (possibly complex) eigenvalues as does $\mathrm{ad}_A : \mathfrak{W}^d \to \mathfrak{W}^d$. By condition ($\mathrm{P}_1$), for $d > 1$ none of the eigenvalues is zero. □

Exercise 7.1. The Jordan decomposition $A = D + N$ for $A \in \mathfrak{t}^u(n,\mathbb{R})$ and the linearity of $\mathrm{ad}_A(X)$ imply $\mathrm{ad}_A = \mathrm{ad}_D + \mathrm{ad}_N$ in the adjoint representation[10] of $\mathfrak{t}^u(n,\mathbb{R})$ on itself. Under what circumstances is this a Jordan decomposition? Is it true that $\mathrm{spec}(A) = \mathrm{spec}(D)$? △

---

[10] See Section B.1.8 for the adjoint representation.

7.2.2 Linearization: Single Vector Fields

The following is a formal-power-series version of Poincaré's theorem on the linearization of vector fields. The notations $\phi^{(d)}, p^{(d)}$, etc., indicate polynomials of degree d. This approach provides a way to symbolically compute a linearizing coordinate transformation ϕ of the form

$$\phi : \mathbb{R}^n \to \mathbb{R}^n; \phi(0) = 0, \ \phi_*(0) = I.$$

The formal power series at 0 for such a mapping ϕ is of the form

$$\phi(x) = Ix + \phi^{(2)}(x) + \phi^{(3)}(x)\ldots, \quad \text{where } \phi^{(k)} \in \mathfrak{V}^k. \tag{7.13}$$

Theorem 7.5. *Let $p(x) := Px + p^{(2)}(x) + p^{(3)}(x) + \cdots \in C^\omega(R^n)$ where P satisfies the condition* (P_1) *and $p^{(k)} \in \mathfrak{V}^k$. Then there exists a formal power series for a linearizing map for $p(x)$ about* 0.

Proof (Livingston [188]). We seek a real-analytic mapping $\phi : \mathbb{R}^n \to \mathbb{R}^n$ with $\phi(0) = 0$, $\phi_*(0) = I$. Necessarily it has the expansion (7.13) and must satisfy $\phi_*(x)p(x) = P\phi(x)$, so

$$(I + \phi_*^{(2)}(x) + \cdots)(Px + p^{(2)}(x) + \cdots) = P(Ix + \phi^{(2)}(x) + \cdots).$$

The terms of degree $k \geq 2$ satisfy

$$\phi_*^k Px - P\phi^{(k)}(x) + \phi_*^{(k-1)}p^{(2)} + \cdots + p^{(k)}(x) = 0.$$

From condition (P_1), ad_P is invertible on $\mathfrak{V}^k$ if $k < 1$, so for $k \geq 2$

$$\phi^{(k)} = -(\mathrm{ad}_P)^{-1}(p^{(k)} + \phi_*^{(2)}p^{(k-1)} + \cdots + \phi_*^{(k-1)}p^{(2)}(x)).$$

□

7.2.3 Linearization of Lie Algebras

The formal power series version of Theorem 7.6 was given by Hermann [124] using some interesting Lie algebra cohomology methods (the Whitehead lemmas, Jacobson [143]) that are applicable to semisimple Lie algebras; they are echoed in the calculations with ad_A in this section. The relevant cohomology groups are all zero.

Convergence of the ϕ series was shown by Guillemin and Sternberg [112]. Independently, Kušnirenko [171] stated and proved Theorem 7.6 in 1967.

Theorem 7.6 (Guillemin–Sternberg–Kušnirenko). *Let $\tilde{\mathfrak{g}}$ be a representation of a finite-dimensional Lie algebra $\mathfrak{g}$ by C^ω vector fields all vanishing at a point p of a C^ω manifold $\mathcal{M}^n$. If $\mathfrak{g}$ is semisimple then $\tilde{\mathfrak{g}}$ is equivalent via a C^ω mapping $\phi : \mathbb{R}^n_* \to \mathcal{M}^n$ to a linear representation $\mathfrak{g}^1$ of $\mathfrak{g}$.*

Remark 7.1. The truth of this statement for *compact* groups was long-known; the proof in [112] uses a compact representation of the semisimple group $\mathbf{G}$ corresponding to $\mathfrak{g}$. On this compact group an averaging with Haar measure provides a vector field g whose representation on the manifold commutes with those in $\tilde{\mathfrak{g}}$ and whose Jacobian matrix at 0 is the identity. In the coordinates in which g is linear, so are all the other vector fields in $\mathfrak{g}$.

Flato et al. [91] extended Theorem 7.6 to nonlinear representations of semisimple Lie groups (and some others) on Banach spaces.

△

Theorem 7.7 (Livingston). *A Lie algebra of real-analytic vector fields on $\mathbb{R}^n$ with common equilibrium point 0 and nonvanishing linear terms can be linearized if and only if it commutes with a Poincaré vector field $p(x)$.*

Proof. By Poincaré's theorem 7.2 there exist coordinates in which $p(x) = Px$. For each $a \in \mathfrak{g}$, $[Px, Ax + a^{(2)}(x) + \cdots] = 0$. That means $\mathrm{ad}_p(a^{(k)}) = 0$ for every $k \geq 2$; from Theorem 7.4 ad_p is one-to-one on $\mathfrak{V}^k$, so $a^k = 0$. □

Lemma 7.2 will be used for special cases in Theorem 7.8. The proof, paraphrased from [238], is an entertaining application of transitivity.

Lemma 7.2. *If a real-analytic vector field f on $\mathbb{R}^n$ with $f(0) = 0$ commutes with every $g \in \mathfrak{g}^1$, a transitive Lie algebra of linear vector fields, then f itself is linear.*

Proof. Fix $\xi \in \mathbb{R}^n$; we will show that the terms of degree $k > 1$ vanish. Since $\mathfrak{g}$ is transitive, given $\mu > 0$ and $\zeta \in \mathbb{R}^n_*$ there exists $C \in \mathfrak{g}$ such that $C\zeta = \mu\zeta$. Let λ be the largest of the real eigenvalues of C, with eigenvector $z \in \mathbb{R}^n_*$. Since the Lie group $\mathbf{G}$ corresponding to $\mathfrak{g}$ is transitive, there exists $Q \in \mathbf{G}$ such that $Qz = \xi$. Let $T = QCQ^{-1}$, then $Tx \in \mathfrak{g}^1$; therefore $[Tx, f(x)] = 0$. Also $T\xi = \lambda\xi$ and λ is the largest real eigenvalue of T. For all $j \in \mathbb{Z}_+$ and all x

$$
\begin{aligned}
0 &= [Tx, f^{(j)}(x)] = f_*^{(j)}(x)Tx - Tf^{(j)}(x). \text{ At } x = \xi,\\
0 &= \lambda f_*^{(j)}(\xi)\xi - Tf^{(j)}(\xi) = \lambda \mathbf{e} f^j(\xi) - Tf^{(j)}(\xi)\\
&= j\lambda f^{(j)}(\xi) - Tf^j(\xi).
\end{aligned}
$$

Thus if $f^{(j)}(\xi) \neq 0$ then $j\lambda$ is a real eigenvalue of T with eigenvector $f^{(j)}(\xi)$; but if $j > 1$ that contradicts the maximality of λ. Therefore $f^{*(j)} \equiv 0$ for $j \geq 2$. □

Theorem 7.8 (Sedwick [237, 238]). *Let $f_1, \ldots, f_m$ be C^ω vector fields on a neighborhood U in a C^ω manifold $\mathcal{M}^n$ all of which vanish at some $p \in U$; denote the linear terms in the power series for f_i at p by $F_i := f_{i*}(p)$. Suppose the vector field Lie algebra $\tilde{\mathfrak{g}} = \{f_1, \ldots, f_m\}_{\mathcal{L}}$ is isomorphic to the matrix Lie algebra*

$\mathfrak{g} := \{F_1, \ldots, F_m\}_{\mathcal{L}}$. *If* $\mathfrak{g}$ *is transitive on* $\mathbb{R}^n_*$ *then the set of linear partial differential equations* $f_i(\phi(z)) = \phi_*(z)F_iz$ *has a* C^ω *solution* $\phi : \mathbb{R}^n_* \to U$ *with* $\phi(0) = p, \phi_*(0) = I$, *giving* C^ω*-equivalence of the control systems*

$$\dot{x} = \sum_{i=1}^{m} u_i f_i(x),\ \dot{z} = \sum_{i=1}^{m} u_i F_i z. \tag{7.14}$$

Proof. This proof combines the methods of [188] and [238]. Refer to Section D.2 in Appendix D; each of the transitive Lie algebras has the decomposition $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{c}$ where $\mathfrak{g}_0$ is semisimple [29, 32, 162] and $\mathfrak{c}$ is the center. If the center is $\{0\}$ choose coordinates such that (by Theorem 7.6) $\tilde{\mathfrak{g}} = \tilde{\mathfrak{g}}_0$ is linear. If the center of $\mathfrak{g}$ is $\mathfrak{c}(n, \mathbb{R})$, $\mathfrak{c}(k, \mathbb{C})$, or $\mathfrak{c}(k, (u+iv)\mathbb{R})$, then $\tilde{\mathfrak{g}}$ contains a Poincaré vector field with which everything in $\mathfrak{g}$ commutes; in the coordinates that linearize the center, $\tilde{\mathfrak{g}}$ is linear.[11] Finally consider the case $\mathfrak{c} = \mathfrak{c}(k, i\mathbb{R})$, which does not satisfy P_2, corresponding to II(2), II(4), and II(5); for these the (noncompact) semisimple parts $\mathfrak{g}_0$ are transitive and in the coordinates in which these are linear, by Lemma 7.2 the center is linear too. □

Finally, one may ask whether a Lie algebra $\tilde{\mathfrak{g}}$ containing a Poincaré vector field p will be linear in the coordinates in which $p(x) = Px$. That is true if p is in the center, but in this example it is not.

Example 7.5 (Livingston).

$$p(x) := \begin{bmatrix} x_1 \\ \sqrt{2}x_2 \\ (3\sqrt{2}-1)x_1 \end{bmatrix}, \quad a(x) := \begin{bmatrix} 0 \\ x_1 \\ x_2^2 \end{bmatrix};$$
$$[p, a] = (1 - \sqrt{2})a, \quad \{p, a\}_{\mathcal{L}} = \operatorname{span}(p, a).$$

The Poincaré vector field p is already linear; the power series in Theorem 7.5 is unique, so a cannot be linearized. △

Theorem 7.9. *Let* $\mathfrak{g}$ *be a Lie algebra of real-analytic vector fields on* $\mathbb{R}^n$ *with common equilibrium point* 0 *and nonvanishing linear terms. If* $\mathfrak{g}$ *contains a vector field* $p(x) = Px + p^2(x) + \cdots$ *satisfying*

(P_1') *the eigenvalues* μ_i *of* P *satisfy no relations of the form*

$$\mu_r + \mu_q = \sum_{j=1}^{n} k_j\mu_j,\ k_j \in \mathbb{Z}_+,\ \sum_{j=1}^{n} k_j > 2; \text{ and}$$

(P_2) *the convex hull of* $\{\mu_1, \ldots, \mu_n\}$ *does not contain the origin,*
then $\mathfrak{g}$ *is linear in the coordinates in which* p *is linear.*

[11] Note that $\mathfrak{spin}(9,1)$ has center either $\{0\}$ or $\mathfrak{c}(16, \mathbb{R})$, so the proof is valid for this new addition to the Boothby list.

For the proof see Livingston [187, 188]. By Levi's Theorem B.5, $\mathfrak{g} = \mathfrak{s} + \mathfrak{r}$ where $\mathfrak{s}$ is semisimple and $\mathfrak{r}$ is the radical of $\mathfrak{g}$. Matrices in $\mathfrak{s}$ will have trace 0 and cannot satisfy (P'_1), so this theorem is of interest if the Poincaré vector field lies in the radical but not in the center.

Example 7.6 (Livingston [187]). The Lie subalgebras of $\mathfrak{gl}(2,\mathbb{R})$ are described in Chapter 2. Nonlinear Lie algebras isomorphic to some of them can be globally linearized on $\mathbb{R}^2_*$ by solving the intertwining equation (7.1) by the power series method of this chapter: $\mathfrak{sl}(2,\mathbb{R})$ because it is simple, $\mathfrak{gl}(2,\mathbb{R})$ and $\boldsymbol{\alpha}(\mathbb{C})$ because they contain I. Another such is $\mathbb{R}I \oplus \mathfrak{n}_2$ where $\mathfrak{n}_2 \subset \mathfrak{gl}(2,\mathbb{R})$ is the non-Abelian Lie algebra with basis $\{E, F\}$ and relation $[E, F] = E$. However, $\mathfrak{n}_2$ itself is more interesting because it is weakly transitive. In well-chosen coordinates

$$E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 1+\lambda & 0 \\ 0 & \lambda \end{bmatrix}.$$

Using Proposition 7.9, [187] shows how to linearize several quadratic representations of $\mathfrak{n}_2$, such as this one.

$$e(x) = \begin{bmatrix} x_2 + \alpha x_2^2 \\ 0 \end{bmatrix}, \; f(x) = \begin{bmatrix} (1+\lambda)x_1 + 2\alpha\lambda x_1 x_2 + \beta x_2^2 \\ \lambda x_2 + \alpha\lambda x_2^2 \end{bmatrix}.$$

If $\lambda \neq 1$ the linearizing map (unique if $\lambda \neq 0$) is

$$\phi(x) = \begin{bmatrix} \frac{x_1}{(1+\alpha x_2)^2} + \frac{\beta}{1-\lambda}\frac{x_2^2}{(1+\alpha x_2)^2} \\ \frac{x_2}{1+\alpha x_2} \end{bmatrix}, \quad \text{for } |x_2| < \frac{1}{|\alpha|}.$$

If $\lambda = 1$ then there is a linearizing map that fails to be analytic at the origin:

$$\phi(x) = \begin{bmatrix} \frac{x_1}{(1+\alpha x_2)^2} - \frac{\beta x_2^2}{(1+\alpha x_2)^2} \ln\left(\frac{|x_2|}{(1+\alpha x_2)^2}\right) \\ \frac{x_2}{1+\alpha x_2} \end{bmatrix}, \; \text{for } |x_2| < \frac{1}{|\alpha|}.$$

Problem 7.1. It was mentioned in Section 7.1 that any bilinear system (7.6) can be disguised with any diffeomorphism $\phi : \mathbb{R}^n_* \to \mathbb{R}^n_*$, $\phi(0) = 0$, to cook up a nonlinear system (7.8) for which the necessary conditions ($\mathrm{H}_1, \mathrm{H}_2$) will be satisfied; but suppose $\mathfrak{g}$ is neither transitive nor semisimple. In that case there is no general recipe to construct ϕ and it would not be unique. △

7.3 Related Work

Sussmann [259] has a theorem that can be interpreted roughly as follows. Let M, N be n dimensional simply connected C^ω manifolds, $\mathfrak{g}$ on M^n, $\tilde{\mathfrak{g}}$ on N^n be isomorphic Lie algebras of vector fields that each has full Lie rank on its manifold. Suppose there can be found points $p \in M^n$, $q \in N^n$ for which the isotropy subalgebras $\mathfrak{g}_p, \tilde{\mathfrak{g}}_q$ are isomorphic; then there exists a real-analytic mapping $\phi : M^n \to N^n$ such that $\tilde{\mathfrak{g}} = \phi_* \mathfrak{g}$. The method of proof is to construct ϕ as a graph in the product manifold $P := M^n \times N^n$. For our problem the vector fields of $\tilde{\mathfrak{g}}$ are linear, $N^n = \mathbb{R}^n$, $n > 2$, and the given isomorphism takes each vector field in $\mathfrak{g}$ to its linear part.

For C^k vector fields, $1 \le k \le \infty$, the linearization problem becomes more technical. The special linear group is particularly interesting. Guillemin and Sternberg [112] gave an example of a C^∞ action of the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$ on $\mathbb{R}^3$ which is not linearizable; in Cairns and Ghys [47] it is noted that this action is not integrable to an $\mathrm{SL}(2, \mathbb{R})$ action, and a modification is found that does integrate to such an action. This work also has the following theorems, and gives some short proofs of the Hermann and Guillemin and Sternberg results for the $\mathfrak{sl}(n, \mathbb{R})$ case. The second one is relevant to discrete-time switching systems; $\mathrm{SL}(n, \mathbb{Z})$ is defined in Example 3.

Theorem 7.10 (Cairns and Ghys, Th. 1.1). *For all $n > 1$ and for all $k \in 1, \dots, \infty$, every C^k-action of $\mathrm{SL}(n, \mathbb{R})$ on $\mathbb{R}^n$ fixing 0 is C^k-linearizable.*

Theorem 7.11 (Cairns–Ghys, Th. 1.2).
(a) *For no values of n and m with $n > m$ are there any non-trivial C^1-actions of $\mathrm{SL}(n, \mathbb{Z})$ on $\mathbb{R}^m_*$.*
(b) *There is a C^1-action of $\mathrm{SL}(3, \mathbb{Z})$ on $\mathbb{R}^8_*$ which is not topologically linearizable.*
(c) *There is a C^∞-action of $\mathrm{SL}(2, \mathbb{Z})$ on $\mathbb{R}^2_*$ which is not linearizable.*
(d) *For all $n > 2$ and $m > 2$, every C^∞-action of $\mathrm{SL}(n, \mathbb{Z})$ on $\mathbb{R}^m_*$ is C^∞-linearizable.*

Chapter 8
Input Structures

Input–output relations for bilinear systems can be represented in several interesting ways, depending on what mathematical structure is imposed on the space of inputs $\mathcal{U}$. In Section 8.1 we return to the topics and notations of Section 1.5.2; locally integrable inputs are given the structure of a topological concatenation semigroup, following Sussmann [261]. In Section 8.2 piecewise continuous (Stieltjes integrable) inputs are used to define iterated integrals for Chen–Fliess series; for bilinear systems these series have a special property called rationality. In Section 8.3 the inputs are stochastic processes with various topologies.

8.1 Concatenation and Matrix Semigroups

In Section 1.5.2 the concatenation semigroup for inputs was introduced. This section looks at a way of giving this semigroup a topology in order to sketch the work of Sussmann [261], especially Theorem 8.1, Sussmann's approximation principal for bilinear systems.

A topological semigroup $\mathfrak{S}_{\star,\tau}$ is a Hausdorff topological space with a semigroup structure in which the mapping $\mathfrak{S}_{\star,\tau} \times \mathfrak{S}_{\star,\tau} \to \mathfrak{S}_{\star,\tau} : (a, b) \mapsto ab$ is jointly continuous in a, b. For example, the set $\mathbb{R}^{n\times n}$ equipped with its usual topology and matrix multiplication is a topological semigroup.

The underlying set $\mathcal{U}$ from which we construct the input semigroup $\mathfrak{S}_{\star,\tau}$ is the space $\mathcal{LI}_\Omega$ of locally integral functions on $\mathbb{R}_+$ whose values are constrained to a set Ω that is convex with nonempty interior. At first we will suppose that Ω is compact.

As in Section 1.5.2 we build $\mathfrak{S}_{\star,\tau}$ as follows. The elements of $\mathfrak{S}_{\star,\tau}$ are pairs; if the support of $u \in \mathcal{LI}_\Omega$ is $\operatorname{supp}(u) := [0, T_u]$ then $\bar{u} := (u, [0, T_u])$. The semigroup operation $\bar{\star}$ is defined by

$$(\bar{u}, \bar{v}) \mapsto \bar{u}\bar{\star}\bar{v} = \{u \star v, [0, T_u + T_v]\}$$

D.L. Elliott, *Bilinear Control Systems*, Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_8, © Springer Science+Business Media B.V. 2009

where the concatenation $(u, v) \to u \star v$ is defined in (1.36).

For $\mathfrak{S}_{\star,\tau}$, we choose a topology $\mathcal{T}_w$ of weak convergence, under which it is compact on finite time intervals. Let $\mathrm{supp}(u^{(k)}) = [0, T_k]$, $\mathrm{supp}(u) = [0, T]$, then we say

$$\bar{u}_k \to_w \bar{u} \text{ in } \mathcal{T} \text{ if } \lim_{k\to\infty} T_k = T \text{ and } u^{(k)} \text{ converges weakly to } u, \text{ i.e.,}$$

$$\lim_{k\to\infty} \int_\alpha^\beta u^{(k)}(t)\,dt = \int_\alpha^\beta u(t)\,dt \text{ for all } \alpha, \beta \text{ such that } 0 \le \alpha \le \beta \le T.$$

Also, as in Section 1.5.2, $\mathbf{S}_\Omega$ is the matrix semigroup of transition matrices; it is convenient to put the initial value problem in integral equation form here. With

$$F(u(t)) := A + \sum_{i=1}^m u_i(s) B_i,$$

$$X(t; u) = I + \int_0^t F(u(s)) X(s; u)\,ds, \;\; u \in \mathcal{LI}_\Omega. \tag{8.1}$$

The set of solutions $X(\cdot;\cdot)$ restricted to $[0, T]$ can be denoted here by $\mathcal{X}$. The following is a specialization of Sussmann [256, Lemma 2] and its proof; it is cited in [261].

Lemma 8.1. *Suppose $\Omega \subset \mathbb{R}^m$ is compact and that $u^{(j)} \in \mathcal{LI}_\Omega$ converges in $\mathcal{T}_w$ to $u \in \mathcal{LI}_\Omega$; then the corresponding solutions of* (8.1) *satisfy $X(t; u^{(j)}) \to X(t; u)$ uniformly on $[0, T]$.*

Proof. Since Ω is bounded, the solutions in $\mathcal{X}$ are uniformly bounded; their derivatives $\dot{X}$ are also uniformly bounded, so $\mathcal{X}$ is equicontinuous. By the Ascoli–Arzela Theorem $\mathcal{X}$ is compact in the topology of uniform convergence. That is, every sequence in $\mathcal{X}$ has a subsequence that converges uniformly to some element of $\mathcal{X}$. To establish the lemma it suffices to show that every subsequence of $\{X(t; u^{(j)})\}$ has a subsequence that converges uniformly to the same $X(t; u)$.

Now suppose that a subsequence $\{v^{(k)}\} \subset \{u^{(j)}\}$ converges in $\mathcal{T}_w$ to the given u and that $X(t; v^{(k)})$ converges uniformly to some $Y(\cdot) \in \mathcal{X}$. From (8.1)

$$X(t; v^{(k)}) = I + \int_0^t F(v^{(k)}(s)[X(s; v^{(k)}) - Y(s)]\,ds + \int_0^t F(v^{(k)}(s))Y(s)\,ds.$$

From the weak convergence $v^{(k)} \to v$ and the uniform convergence $X(t; v^{(k)}) \to Y(\cdot)$,

$$Y(t) = I + \int_0^t F(v^{(k)}(s) Y(s\,ds; \quad \text{so } Y(t) \equiv X(t; u).$$

□

From Lemma 8.1 the mapping $\mathfrak{S}_{\star,\tau} \to \mathbf{S}_\Omega$ is a *continuous* semigroup representation.

For continuity of this representation in $\mathcal{T}_w$ the assumption that Ω is bounded is essential; otherwise limits of inputs in $\mathcal{T}_w$ may have to be interpreted as generalized inputs; they are recognizable as Schwarz distributions when $m = 1$ but not for $m \geq 2$. Sussmann [261, §5] gives two examples for $\Omega := \mathbb{R}^m$. In the case $m = 1$ let $\dot{X} = X + uJX$, $X(0) = I$. For $k \in \mathbb{N}$ let $u^{(k)}(t) = k,\ 0 \leq t \leq \pi/k$, $u^{(k)}(t) = 0$ otherwise, then $X(1; u^{(k)}(t) \to \exp(\pi)I = -I$ as $u^{(k)}(t) \to \delta(t)$ (the "delta-function" Schwarz distribution). The second example has two inputs.

Example 8.1. On $[0,1]$ define the sequence[1] of pairs of $\mathcal{PK}$ functions $\{w^{(k)}\} := \{(u^{(k)}, v^{(k)}), k \in \mathbb{N}\}$ defined with $0 \leq j \leq k$ by

$$u^{(k)}(t) = \begin{cases} \sqrt{k}, & \frac{4j}{4k} \leq t \leq \frac{4j+1}{4k} \\ -\sqrt{k}, & \frac{4j+2}{4k} \leq t \leq \frac{4j+3}{4k} \\ 0, & \text{otherwise} \end{cases}, \quad v^{(k)}(t) = \begin{cases} \sqrt{k}, & \frac{4j+1}{4k} \leq t \leq \frac{4j+2}{4k} \\ -\sqrt{k}, & \frac{4j+3}{4k} \leq t \leq \frac{4j+4}{4k} \\ 0, & \text{otherwise} \end{cases}.$$

It is easily seen that $\int_0^1 w^{(k)}(t)\phi(t)\,dt \to 0$ for all $\phi \in C^1[0,1]$, so $w^{(k)} \to 0$ in $\mathcal{T}_w$; that is, in the space of Schwarz distributions $\{w^{(k)}\}$ is a null sequence. If we use the pairs $(u^{(k)}, v^{(k)})$ as inputs in a bilinear system $\dot{X} = u^{(k)}AX + v^{(k)}BX$, $X(0) = I$ and assume $[A,B] \neq 0$ then as $k \to \infty$ we have $X(1; w^{(k)}) \to \exp([A,B]/4) \neq I$, so this representation cannot be continuous with respect to $\mathcal{T}_w$. △

The following approximation theorem, converse to Lemma 8.1, refers to input–output mappings for observed bilinear systems with $u := (u_1, \ldots, u_m)$ is bounded by μ:

$$\dot{x} = Ax + \sum_{i=1}^{m} u_i B_i x, \ \ \psi_u(t) = c^\top x(t),$$
$$\sup\{|u_i(t)|, 0 \leq t \leq T, i \in 1, \ldots, m\} \leq \mu.$$

Theorem 8.1 (Sussmann [261], Th. 3). *Suppose that ϕ is an arbitrary function which to every input u defined on $[0,T]$ and bounded by μ assigns a curve $\phi_u : [0,T] \to \mathbb{R}$; that ϕ is causal; and that it is continuous in the sense of Lemma 8.1. Then for every $\epsilon > 0$ there is a bilinear system whose corresponding input–output mapping ψ satisfies $|\phi_u(t) - \psi_u(t)| < \epsilon$ for all $t \in [0,T]$ and all $u : [0,T] \to \mathbb{R}^m$ which are measurable and bounded by μ.*

Sussmann [260] constructed an interesting topology on $\mathcal{LI}^m_{\mathbb{R}}$: this topology $\mathcal{T}_S$ is the weakest for which the representation $\mathfrak{S}_{\star,\tau} \to \mathbf{S}_\Omega$ is continuous. In $\mathcal{T}_S$ the sequence $\{w^{(k)}\}$ in Example 8.1 is convergent to a nonzero generalized input $w^{(\infty)}$. Sussmann [261, 263] shows how to construct a Wiener process as such a generalized input; see Section 8.3.3.

[1] Compare the control (u_1, u_2) used in Example 2.1; here it is repeated k times.

8.2 Formal Power Series for Bilinear Systems

For the theory of formal power series (f.p.s.) in commuting variables see Henrici [121]. F.p.s in noncommuting variables are studied in the theory of formal languages; their connection to nonlinear system theory was first seen by Fliess [92]; for an extensive account see Fliess [93]. Our notation differs, since his work used nonlinear vector fields.

8.2.1 Background

We are given a free concatenation semigroup $\mathfrak{A}*$ generated by a finite set of noncommutative indeterminates $\mathfrak{A}$ (an alphabet). An element $\mathbf{w} = a_1 a_2 \dots a_k \in \mathfrak{A}*$ is called a word; the *length* $|\mathbf{w}|$ of this example is k. The identity (empty word) in $\mathfrak{A}*$ is called ι: $\iota\mathbf{w} = \mathbf{w}\iota = \mathbf{w}$, $|\iota| = 0$. For repeated words we write $\mathbf{w}^0 := \iota$, $a^2 := aa$, $\mathbf{w}^2 := \mathbf{ww}$, and so on. For example use the alphabet $\mathfrak{A} = \{a, b\}$; the words of $\mathfrak{A}*$ (ordered by length and lexicographic order) are $\iota, a, b, a^2, ab, ba, b^2$, etc. Given a commutative ring $\mathbf{R}$ one can talk about a noncommutative polynomial $p \in \mathbf{R}{<}\mathfrak{A}{>}$ (the free $\mathbf{R}$-algebra generated by $\mathfrak{A}$) and a formal power series $S \in \mathbf{R}{\ll}\mathfrak{A}{\gg}$ with coefficients $S_\mathbf{w} \in \mathbf{R}$,

$$S = \sum \{S_\mathbf{w}\mathbf{w} |\, \mathbf{w} \in \mathcal{A}*\}.$$

Given two f.p.s. S, S' one can define addition, (Cauchy) product[2], and the multiplicative inverse S^{-1} which exists if and only if S has a nonzero constant term S_ι.

$$S + S' := \sum \{(S_\mathbf{w} + S'_\mathbf{w})\mathbf{w} |\, \mathbf{w} \in \mathfrak{A}*\},$$
$$SS' := \sum \Big\{\Big(\sum_{\mathbf{vv}'=\mathbf{w}} S_\mathbf{v} S'_{\mathbf{v}'}\Big)\mathbf{w} \,\Big|\, \mathbf{w} \in \mathfrak{A}*\Big\}.$$

Example: Let $S_1 := 1 + \alpha_1 a + \alpha_2 b,\ \alpha_i \in \mathbb{F};$ then
$S_1^{-1} = 1 - \alpha_1 a - \alpha_2 b + \alpha_1^2 a^2 + \alpha_1\alpha_2(ab + ba) + \alpha_2^2 b^2 + \cdots$

The smallest subring $\mathbf{R}{<}(\mathfrak{A}){>} \subset \mathbf{R}{\ll}\mathfrak{A}{\gg}$ that contains the polynomials (elements of finite length) and is closed under rational operations is called the ring of rational f.p.s. Fliess [93] showed that as a consequence of the Kleene–Schützenberger Theorem, a series $S \in \mathbf{R}{\ll}\mathfrak{A}{\gg}$ is rational if and only if there exist an integer $n \geq 1$, a representation $\rho : \mathfrak{A}* \to \mathbf{R}^{n\times n}$ and vectors $c, b \in \mathbf{R}^n$ such that

[2] The Cauchy, Hadamard and shuffle products of f.p.s. are discussed and applied by Fliess [92, 93].

$$S = \sum_{\mathbf{w} \in \mathfrak{A}*} c^\top \rho(\mathbf{w}) b.$$

An application of these ideas in [93] is to problems in which the alphabet is a set of C^ω vector fields $\mathfrak{A} := \{\mathbf{a}, \mathbf{b}_1, \ldots, \mathbf{b}_m\}$. If the vector fields are linear the representation is defined by $\rho(\mathfrak{A}) = \{A, B_1, \ldots, B_m\}$, our usual set of matrices; thus

$$\rho\left(\sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbf{a}^k\right) = \exp(tA).$$

8.2.2 Iterated Integrals

Let $\mathfrak{A} := \{a_0, a_1, \ldots, a_m\}$ and $\mathcal{U} := \mathcal{PC}[0, T]$, the piecewise continuous real functions on $[0, T]$; integrals will be Stieltjes integrals. To every nonempty word $\mathbf{w} = a_{j_n} \ldots a_{j_0}$ of length $n = |\mathbf{w}|$ will be attached an iterated integral[3]

$\int_0^t dv_{j_1} \ldots dv_{j_k}$ defined by recursion on k; for $i \in 1, \ldots, m,\ t \in [0, T],\ u \in \mathcal{U}$

$$v_0(t) := t,\ v_i(t) := \int_0^t u_i(\tau)\, d\tau;\ \int_0^t dv_i := v_i(t);$$
$$\int_0^t dv_{j_k} \ldots dv_{j_0} := \int_0^t dv_{j_k} \int_0^\tau dv_{j_{k-1}} \ldots dv_{j_0}.$$

Here is a simple example:

$$\mathbf{w} = a_0^n a_1, \quad S_{\mathbf{w}}(t; u)\mathbf{w} = \int_0^t \frac{(t-\tau)^n}{n!} u_1(\tau)\, d\tau\, a_0^n a_1.$$

Fliess [93, II.4] gives a bilinear example obtained by iterative solution of an integral equation (compare (1.33))

$$\dot{x} = B_0 x + \sum_{i=1}^{m} u_i(t) B_i x,\ x(0) = \xi,\ y(t) = c^\top x;$$
$$y(t) = c^\top \left(I + \sum_{k \geq 0} \sum_{j_0, \ldots, j_k = 0}^{k} B_{j_k} \ldots B_{j_0} \int_0^t dv_{j_k} \ldots dv_{j_0} \right) \xi. \tag{8.2}$$

The Chen–Fliess series (8.2) converges on $[0, T]$ for all $u \in \mathcal{U}$. Fliess [93, Th. II.5] provides an approximation theorem that can be restated for our

[3] See Fliess [92, 93], applying the iterated integrals of Chen [53, 55].

purposes (compare [261]). Its proof is an application of the Stone–Weierstrass Theorem.

Theorem 8.2. *Let $C \subset C^0([0,T])^m \to \mathbb{R}$ be compact in the topology $\mathcal{T}_u$ of uniform convergence, $\phi : C \to \mathbb{R}$ a causal continuous functional $[0,T] \times C \to \mathbb{R}$; in $\mathcal{T}_u$, ϕ can be arbitrarily approximated by a causal rational or polynomial functional* (8.2).

Approximation of causal functional series by rational formal power series permits the approximation of the input–output mapping of a nonlinear control system by a bilinear control system with linear output. The dimension n of such an approximation may be very large.

Sussmann [266] considered a formal system

$$\dot{X} = X \sum_{i=1}^{m} u_i B_i, \;\; X(0) = I,$$

where the B_i are noncommuting indeterminates and I is the formal identity. It has a Chen series solution like (8.2), but instead, the Theorem of [266] presents a formal solution for (8.2.2) as an infinite product[4]

$$\prod_{0}^{\infty} \exp(V_i(t)A_i),$$

where $\{A_i, i \in \mathbb{N}\}$ is (see Section B.1.6) a Philip Hall basis for the free Lie algebra generated by the indeterminates $\{B_1, \ldots, B_m\}$, exp is the formal series for the exponential and the $V_i(t)$ are iterated integrals of the controls. This representation, mentioned also on page 75, then can be applied by replacing the B_i with matrices, in which case [266] shows that the infinite product converges for bounded controls and small t. The techniques of this paper are illuminating. If $\{B_1, \ldots, B_m\}_{\mathcal{L}}$ is nilpotent it can be given a finite Philip Hall basis so that the representation is a finite product. This product representation has been applied by Lafferriere and Sussmann [172] and Margaliot [199].

Remark 8.1 (Shuffle product). Fliess [93] and many other works on formal series make use of the shuffle[5] binary operator in $\mathbf{R} < \mathfrak{A} >$ or $\mathbf{R} \ll \mathfrak{A} \gg$; it is characterized by $\mathbf{R}$-linearity and three identities:

$$\text{For all } a, b \in \mathfrak{A},\; \mathbf{w}, \mathbf{w}' \in \mathfrak{A}*, \quad \iota \sqcup\!\sqcup\, a = a, a \sqcup\!\sqcup\, \iota = a,$$
$$(a\mathbf{w}) \sqcup\!\sqcup\, (b\mathbf{w}') = a(\mathbf{w} \sqcup\!\sqcup\, (b\mathbf{w}')) + b((a\mathbf{w}) \sqcup\!\sqcup\, \mathbf{w}');$$

thus $a \sqcup\!\sqcup\, b = ab + ba$ and $(ab) \sqcup\!\sqcup\, (cd) = abcd + acbd + acdb + cdab + cadb + cabd.$

[4] The left action in (8.2.2) makes possible the left-to-right order in the infinite product; compare (1.30).

[5] The $\sqcup\!\sqcup$ product is also called the Hurwitz product; in French it is *le produit melange*.

It is pointed out in [93] that given two input–output mappings $u \mapsto y(t)$ and $u \mapsto z(t)$ given as rational (8.2), their shuffle product $y ⧢ z$ is a rational series that represents the product $u \mapsto y(t)z(t)$ of the two input–output maps. See Section 5.1 for a bilinear system (5.7) that realizes this mapping. △

8.3 Stochastic Bilinear Systems

A few definitions and facts from probability theory are summarized at the beginning of Section 8.3.1. Section 8.3.2 introduces randomly switched systems, which are the subject of much current writing. Probability theory is more important for Section 8.3.3's account of diffusions; bilinear diffusions and stochastic systems on Lie groups need their own book.

8.3.1 Probability and Random Processes

A probability space is a triple $\mathfrak{P} := (\boldsymbol{\Omega}, \mathcal{A}, \Pr)$. Here $\boldsymbol{\Omega}$ is (w.l.o.g.) the unit interval $[0, 1]$; $\mathcal{A}$ is a collection of measurable subsets $E \subset \boldsymbol{\Omega}$ (such a set E is called an event) that is closed under intersection and countable union; $\mathcal{A}$ is called a sigma-algebra.[6] A probability measure Pr on $\mathcal{A}$ must satisfy Kolmogorov's three axioms: for any event E and countable set of disjoint events E_1, E_2, etc.

$$\Pr(E) \geq 0,\ \Pr(\boldsymbol{\Omega}) = 1,\ \Pr\left(\bigcup\nolimits_i E_i\right) = \sum_i \Pr(E_i).$$

Two events E_1, E_2 such that $\Pr(E_1 \cup E_2) = \Pr(E_1)\Pr(E_2)$ are called stochastically independent events. If $\Pr(E) = 1$ one writes "E w.p. 1".

A real random variable $f(\omega)$ is a real function on $\boldsymbol{\Omega}$ such that for each α $\{\omega | f(\omega \leq \alpha\} \in \mathcal{A}$; we say that f is $\mathcal{A}$-measurable. The expectation of f is the linear functional $\mathrm{E}f := \int_\Omega f(\omega)\ \Pr(d\omega)$; if Ω is discrete, $\mathrm{E}f := \sum_i f(\omega_i)\Pr(\omega_i)$.

Analogously we have random vectors $f = (f_1, \dots, f_n)$, and a vector stochastic process on $\mathfrak{P}$ is an $\mathcal{A}$-measurable function $v : \mathbb{R} \times \Omega \to \mathbb{R}^m$ (assume for now its sample functions $v(t; \omega)$ are locally integrable w.p. 1). Often $v(t; \omega)$ is written $v(t)$ when no confusion should result.

Given a sigma-algebra $\mathcal{B} \subset \mathcal{A}$ with $\Pr(\mathcal{B}) > 0$ define the conditional probability

[6] The symbol $\boldsymbol{\Omega}$ is in boldface to distinguish it from the Ω used for a constraint set. The sigma-algebra $\mathcal{A}$ may as well be assumed to be the set of Lebesgue measurable sets in $[0, 1]$.

$$\Pr[E|\mathcal{B}] := \Pr(A \cup \mathcal{B})/\Pr(\mathcal{B}) \text{ and } \mathrm{E}[f|\mathcal{B}] := \int_{\Omega} f(\omega) \Pr(d\omega|\mathcal{B}),$$

the conditional expectation of f given $\mathcal{B}$. The theory of stochastic processes is concerned with *information*, typically represented by some given family of nested sigma-algebras[7]

$$\{\mathcal{A}_t, t \geq 0 | \mathcal{A}_t \subset \mathcal{A}_{t+\tau} \text{ for all } \tau > 0\}.$$

If each value $v(t)$ of a stochastic process is stochastically independent of $\mathcal{A}_T$ for all $T > t$ (the future) then the process is called non-anticipative. Nested sigma-algebras are also called *classical information structures*; $\mathcal{A}_t$ contains the events one might know at time t if he/she has neither memory loss nor a time machine. We shall avoid the technical conditions that this theory requires as much as possible.

Definition 8.1. By a stochastic bilinear system is meant

$$\dot{x} = Ax + \sum_{i=1}^{m} v_i(t;\omega)B_i x, \ x \in \mathbb{R}^n, \ t \in \mathbb{R}_+, \ x(0) = \xi(\omega), \tag{8.3}$$

where on probability space $\mathfrak{P}$ the initial condition ξ is $\mathcal{A}_0$-measurable and the m-component process v induces the nested sigma-algebras: $\mathcal{A}_t$ contains $\mathcal{A}_0$ and the events defined by $\{v(s), 0 < s \leq t\}$. △

The usual existence theory for linear differential equations, done for each ω, provides random trajectories for (8.3). The appropriate notions of stability of the origin are straightforward in this case; one of them is certain global asymptotic stability. Suppose we know that A, B are Hurwitz and $\{A, B\}_{\mathcal{L}}$ is solvable; the method of Section 6.3.3 provides a quadratic Lyapunov function $V \gg 0$ such that $\Pr\{V(x_t) \downarrow 0\} = 1$ for (8.3), from which global asymptotic stability holds w.p.1.

8.3.2 Randomly Switched Systems

This section is related to Section 6.3.3 on arbitrarily switched systems; see Definitions 6.4, 6.5 and Propositions 6.5, 6.6. Bilinear systems with random inputs are treated in Pinsky's monograph [218].

Definition 8.2 (Markov chain). Let the set $\mathcal{N}_\nu \in 1, \dots, \nu$ be the state space for a Markov chain $\{v_t \in \mathcal{N}_\nu | t \geq 0\}$ constructed as follows from a probability

[7] Nested sigma-algebras have been called classical information patterns; knowledge of Pr is irrelevant. The Chinese Remainder Theorem is a good source of examples. Consider the roll of a die, fair or unfair: $\Omega := \{1, 2, \dots, 6\}$. Suppose the observations at times 1, 2 are $y_1 = \omega \mod 2$, $y_2 = \omega \mod 3$. Then $\mathcal{A}_0 := \{\emptyset\}$; $\mathcal{A}_1 := \{\{1, 3, 5\}, \{2, 4, 6\}\}$; $\mathcal{A}_2 := \{\Omega\}$.

space $\mathfrak{P}$, a positive number λ, and a $\nu \times \nu$ rate matrix $G := \{g_{i,j}\}$ that satisfies $\sum_1^\nu g_{i,j} = 0$ and $g_{i,i} = -\sum_{j \neq i} g_{i,j}$. The state $v(t)$ is constant on time intervals (t_k, t_{k+1}) whose lengths $\tau_k := t_{k+1} - t_k$ are stochastically independent random variables each with the exponential distribution $\Pr\{\tau_i > t\} = \exp(-\lambda t)$. The transitions between states $v(t)$ are given by a transition matrix $P(t)$ whose entries are

$$p_{i,j}(t) = \Pr\{v_t = j \,|\, v(0) = i\}$$

that satisfies the Chapman–Kolmogorov equation $\dot{P} = GP$ with $P(0) = I$. Since we can always restrict to an invariant subset of $\mathcal{N}_\nu$ if necessary, there is no loss in assuming that G is irreducible; then the Markov chain has a unique stationary distribution $\mu := (\mu_1, \ldots, \mu_\nu)$ satisfying $P(t)\mu = \mu$ for all $t \geq 0$. Let $\mathbf{E}_\mu$ denote the expectation with respect to μ. △

Pinsky [218] considers bilinear systems $\dot{x} = Ax + \phi(v_t)Bx$ whose scalar input $\phi(v_t)$ is stationary with zero mean, $\sum_{i=1}^\nu \phi(i)\mu_i = 0$, as in the following example. If ϕ is bounded, the rate of growth of the trajectories is estimated by the Lyapunov spectrum defined in Section 1.7.2.

Example 8.2. Consider a Markov chain $v_t \in \mathcal{N}_2 := \{1, 2\}$ such that for $\lambda_1 > 0$ and $\lambda_2 > 0$

$$G = \begin{bmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{bmatrix}; \ \lim_{t\to\infty} P(t) = \begin{bmatrix} \lambda_2 & \lambda_1 \\ \lambda_2 & \lambda_1 \end{bmatrix}, \ \mu = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix};$$

and a bilinear system on $\mathbb{R}^n$ driven by $\phi(v_t)$,

$$\dot{x} = A + \phi(v_t)B, \ \phi(1) = -1, \phi(2) = 1, \tag{8.4}$$

whose trajectories are $x(t) = X(t; v)\xi$ where X is the transition matrix. In the special case $[A, B] = 0$ we can use Example 1.14 to find the largest Lyapunov exponent. Otherwise, Pinsky [218] suggests the following method. Let $V : \mathbb{R}^n \to \mathbb{R}^2$ be continuously differentiable in x, and as usual take $\mathbf{a} = x^\top A^\top \partial/\partial x$, $\mathbf{b} = x^\top B^\top \partial/\partial x$. The resulting process on $\mathbb{R}^n \times \mathcal{N}_\nu$ is described by this Kolmogorov backward equation

$$\begin{cases} \dfrac{\partial V_1}{\partial t} & = \mathbf{a}V_1(x) - \mathbf{b}V_1(x) - \lambda_1 V_1(x) + \lambda_1 V_2(x) \\ \dfrac{\partial V_2}{\partial t} & = \mathbf{a}V_2(x) + \mathbf{b}V_2(x) + \lambda_2 V_1(x) - \lambda_2 V_2(x) \end{cases} \tag{8.5}$$

which is a second-order partial differential equation of hyperbolic type. Compare the generalized telegrapher's equation in Pinsky [218]; there estimates for the Lyapunov exponents of (8.4) are obtained from an eigenvalue problem for (8.5). △

8.3.3 Diffusions: Single Input

Our definition of a diffusion process is a stochastic process $x(t)$ with continuous trajectories that satisfies the strong Markov property — conditioned on a stopping time $\boldsymbol{t}(\omega)$, future values are stochastically independent of the past. The book by Ikeda and Watanabe [138] is recommended, particularly because of its treatment of stochastic integrals and differential equations of both Itō and Fisk–Stratonovich types. There are other types of diffusions, involving discontinuous (jump) processes, but they will not be dealt with here.

The class of processes we will discuss is exemplified by the standard Wiener process (also called a Brownian motion): $\{w(t;\omega), t \in \mathbb{R}_+, \omega \in \boldsymbol{\Omega}\}$ (often the ω argument is omitted) is a diffusion process with $w(0) = 0$ and $\mathrm{E}[w(t)] = 0$ such that differences $w(t+\tau) - w(t)$ are Gaussian with $\mathrm{E}(w(t+\tau) - w(t))^2 = |\tau|$ and differences over non-overlapping time intervals are stochastically independent. The Wiener process induces a nested family of sigma-algebras $\{\mathcal{A}_t, t \geq 0\}$; $\mathcal{A}_0$ contains any events having to do with initial conditions, and for larger times $\mathcal{A}_t$ contains $\mathcal{A}_0$ and the events defined on Ω by $\{w(s,\omega), 0 < s \leq t\}$.

Following [138, p. 48], if $f : (\mathbb{R}_+, \boldsymbol{\Omega}) \to \mathbb{R}$ is a random step function constant on intervals (t_i, t_{i+1}) of length bounded by $\delta > 0$ on $\mathbb{R}_+$ and $f(t;\cdot)$ is $\mathcal{A}_t$-measurable, then define its indefinite Itō stochastic integral with respect to w as

$$\mathrm{I}(f,t;\omega) = \int_0^t f(s,\omega)\,dw(s;\omega) \quad \text{given on } t_n \leq t \leq t_{n+1} \text{ by}$$

$$f(t_n;\omega)(w(t;\omega) - w(t_n;\omega)) + \sum_{i=0}^{n-1} f(t_i;\omega)(w(t_{i+1};\omega) - w(t_i;\omega)),$$

$$\text{or succinctly,} \quad \mathrm{I}(f,t) := \sum_{i=0}^{\infty} f(t_i)(w(t \wedge t_{i+1}) - w((t \wedge t_i)). \tag{8.6}$$

It is is straightforward to extend this definition to any general non-anticipative square-integrable f by taking limits in L_2 as $\delta \downarrow 0$; see Clark [60] or [138, Ch. 1]. The Itō integral has the martingale property

$$\mathrm{E}[\mathrm{I}(f,t)|\mathcal{A}_s] = \mathrm{I}(f,s) \text{ w.p.1}, s \leq t.$$

By the *solution* of an Itō stochastic differential equation such as $dx = Ax\,dw(t)$ is meant the solution of the corresponding Itō integral equation $x(t) = x(0) + \mathrm{I}(Ax,t)$. The Itō calculus has two basic rules: $\mathrm{E}dw(t) = 0$ and $\mathrm{E}dw^2(t) = dt$. They suffice for the following example.

Example 8.3. On $\mathbb{R}^2_*$ let $dx(t) = Jx(t)\,dw(t)$ with $J^2 = -I$, $J^\top J = I$, and $x(0) = \xi$; then if a sample path $w(t;\omega)$ is of bounded variation (which has probability zero) $x^\top(t)x(t) = \xi^\top\xi$ and the solution for this ω lives on a circle. For the Itō solution, $\mathrm{E}[d(x^\top(t)x(t))|\mathcal{A}_t] = \left(\mathrm{E}[x^\top(t)x(t)\middle|\mathcal{A}_t]\right)dt > 0$, so the expected norm of this solution grows with time; this is closely related to a construction of the Itō solution using the Euler discretization (1.52) of Section 1.8. As in Section 4.5 the transition matrices fail to lie in the Lie group one might hope for. $\triangle$

For Fisk–Stratonovich (F–S) integrals and stochastic differential equations from the viewpoint of applications see Stratonovich [253], and especially Kallianpur and Striebel [153]. They are more closely related to control systems, especially bilinear systems, but do not have the martingale property. Again, suppose f is a step function constant on intervals (t_i, t_{i+1}) of length bounded by δ and $f(t;\omega)$ is non-anticipative then following [138, Ch. III] an indefinite F–S stochastic integral of f with respect to w on $\mathbb{R}_+$ is

$$\mathrm{FS}(f,t) = \int_0^t f(s) \circ dw(s) := \sum_{i=0}^{n} f\left(\frac{t_i + t_{i+1}}{2}\right)(w(t_{i+1}) - w(t_i)) \tag{8.7}$$

and taking limits in probability as $\delta \downarrow 0$ it can be extended to more general non-anticipative integrands without trouble. The symbol $\circ$ is traditionally used here to signify that this is a *balanced* integral (not composition). The F–S integral as defined in [138] permits integrals $\mathrm{FS}(f,t;v) = \int_0^t f(s) \circ dv(s)$ where $v(t;\omega)$ is a quasimartingale;[8] the quasimartingale property is inherited by $\mathrm{FS}(f,t;v)$. For F—S calculus the chain rule for differentiation is given in Ikeda and Watanabe [138, Ch. 3, Th. 1.3] as follows.

Theorem 8.3. *If $f_1, \ldots, f_m$ are quasimartingales and $g : \mathbb{R}^m \to \mathbb{R}$ is a C^3 function then*

$$dg(f_1,\ldots,f_m) = \sum_1^m \frac{\partial g}{\partial f_i} \circ df_i.$$

Clark [60] showed how F–S stochastic differential equations on differentiable manifolds can be defined. For their relationship with nonlinear control systems see Elliott [81, 82], Sussmann [264], and Arnold and Kliemann [9].

To the single input system $\dot{x} = Ax + uBx$ with random ξ corresponds the "noise driven" bilinear stochastic differential equation of F–S type

$$dx(t) = Ax\,dt + Bx \circ dw(t),\ x(0) = \xi \ \text{ or} \tag{8.8}$$
$$x(t) = \xi + \int_0^t Ax(s)\,ds + \int Bx(s) \circ dw(s),$$

where each $dw(t)$ is an increment of a Wiener process $\mathfrak{W}^1$. The way these differ from Itō equations is best shown by their solutions as limits of discretizations. Assume that $x(0) = \xi \in \mathbb{R}^n_*$ is a random variable with zero mean. To approximate sample paths of the F–S solution use the midpoint method of (1.54) as in the following illustrative example.

[8] A quasimartingale (Itō's term for a continuous semimartingale) is the sum of a martingale and a $\mathcal{A}_t$-adapted process of bounded variation; see Ikeda and Watanabe [138].

Example 8.4. Consider the F–S stochastic model of a rotation ($J^2 = -I, J^\top J = I$): $dx = Jx \circ dw(t)$ on $[0, T]$. Suppose that for $k = 0, 1, \ldots, t_N = T$ our νth midpoint discretization with $\max(t_{k+1} - t_k) \le \delta/\nu$ and $\Delta w_k := w(t_{k+1} - w(t_k)$ is

$$x^\nu(t_{k+1}) = x^\nu(t_k) + J\frac{x^\nu(t_k) + x^\nu(t_{k+1})}{2}\Delta w_k, \text{ so}$$
$$x^\nu(t_{k+1}) = F_k^\nu(t_k) \text{ where}$$
$$F_k := (I + \frac{1}{2}\Delta w_k J)(I - \frac{1}{2}\Delta w_k J)^{-1}.$$

The matrix F_k is orthogonal and the sequence $\{x^\nu(k)\}$ is easily seen to be a Markov process that lives on a circle; the limit stochastic process inherits those properties. △

In solving an ordinary differential equation on $\mathbb{R}^n$ the discretization, whether Euler or midpoint, converges to a curve in $\mathbb{R}^n$, and the solutions of a differential equation on a smooth manifold $\mathcal{M}^n$ can be extended from chart to chart to live on $\mathcal{M}^n$. We have seen that the Itō method, with its many advantages, lacks that property; but the Fisk–Stratonovich method has it. As shown in detail in [60] and in [138, Ch. V], using the chain rule, a local solution can be extended from chart α to overlapping chart β by a differentiable mapping $\phi_{\alpha\beta}$ on the chart overlap.

Another way to reach the same conclusion is that the Kolmogorov backward operator (differential generator) associated with the diffusion of (8.8) is $K := \mathbf{a} + \frac{1}{2}\mathbf{b}^2 a$ (as in [82]), which is tensorial. If a smooth ϕ is an invariant for $\dot{x} = Ax + uBx$ we know that $\mathbf{a}\phi = 0$ and $\mathbf{b}\phi = 0$. Therefore $K\phi = 0$ and by Itō's Lemma $d\phi = K\phi dt + \mathbf{b}\phi dw = 0$, so ϕ is constant along sample trajectories w.p. 1.

Remark 8.2 [81, 82]. By Hörmander's Theorem [130][9] if the Lie algebra $\{\mathbf{a}, \mathbf{b}\}_\mathcal{L}$ is transitive on a manifold then the operator K is hypoelliptic there: if d is a Schwarz distribution, then a solution of $K\phi = d$ is smooth except on the singular support of d. We conclude that if $\{A, B\}_\mathcal{L}$ is transitive K and its formal dual K^* (the Fokker–Plank operator) are hypoelliptic on $\mathbb{R}^n_*$. The transition densities with initial $p(x, 0, \xi) = \delta(t)\delta(x - \xi)$ have support on the closure of the attainable set $\mathcal{A}(\xi)$ for the control system $\dot{x} = Ax + uBx$ with unconstrained controls. If there exists a Lyapunov function $V \gg 0$ such that $KV < 0$ then there exists an invariant probability density $\rho(x)$ on $\mathbb{R}^n_*$. △

Sussmann in a succession of articles has examined stochastic differential equations from a topological standpoint. Referring to [264, Th. 2], Ax satisfies its linear growth condition and $D(Bx) = B$ satisfies its uniform boundedness condition; so there exists a solution $x(t) = X(t; w)\xi$ of (8.8) for each continuous sample path $w(t) \in \mathfrak{W}^1$, depending *continuously* on w. The $\{x(t), t \ge 0,\}$ is a

[9] The discussion of Hörmander's theorem on second-order differential operators in Oleinik and Radkevic [216] is useful.

stochastic process on the probability space $\mathfrak{P}$ equipped with a nested family of sigma-algebras $\{\mathcal{A}_t | \, t \geq 0\}$ generated by $\mathfrak{W}^1$.

For systems whose inputs are in $\mathfrak{W}^m$, $m > 1$, one cannot conclude continuity of the solution with respect to $\boldsymbol{w}$ under its product topology; this peculiarity is the topic of Section 8.3.4.

Exercise 8.1. Use Example 8.4 and Example 1.15 to show that if $M^k \subset \mathbb{R}^n$ is the intersection of quadric hypersurfaces that are invariant for the linear vector fields **a** and **b** then M^k will be invariant for sample paths of $dx = Ax\,dt + Bx \circ dw$.

8.3.4 Multi-Input Diffusions

Let $\mathfrak{W}^1$ denote the space $C[0, T]$ (real continuous functions w on $[0, T]$) with the uniform norm $\|fw\| = \sup_{[0,T]} |w(t)|$, $w(0) = 0$, and a stochastic structure such that differences $w(t + \tau) - w(t)$ are Gaussian with mean 0 and variance $|\tau|$ and differences over non-overlapping time intervals are stochastically independent. The product topology on the space $\mathfrak{W}^m$ of continuous curves $\boldsymbol{w} : [0, T] \to \mathbb{R}^m$ is induced by the norm $\|\boldsymbol{w}\| := \|w_1\| + \cdots + \|w_m\|$. Sussmann [264] used Example 8.1 to show that for $m > 1$ the F–S solution of

$$dx(t) = Ax(t)dt + \sum_{i=}^{m} B_i x(t) \circ dw_i(t) \tag{8.9}$$

is not continuous with respect to this topology on $\mathfrak{W}^m$.

The differential generator K for the Markov process (8.9) is defined for smooth ϕ by

$$K\phi := \left(\mathbf{a} + \frac{1}{2}\sum_{i=1}^{m} \mathbf{b}_i^2\right)\phi = x^\top A^\top \frac{\partial \phi}{\partial x} + \frac{1}{2}\sum_{i=1}^{m}\left(x^\top B_i^\top \frac{\partial}{\partial x}\right)^2 \phi.$$

If $\{\mathbf{a}, \mathbf{b}_1, \ldots, \mathbf{b}_m\}_{\mathcal{L}}$ is transitive on $\mathbb{R}^n_*$ then the operator K is hypoelliptic there (see Remark 8.2). If $m = n$ the operator K is elliptic (shares the properties of the Laplace operator) on $\mathbb{R}^n_*$ if

$$\sum_{i=1}^{n} B_i^\top B_i \gg 0.$$

This page intentionally left blank

Appendix A
Matrix Algebra

A.1 Definitions

In this book the set of integers is denoted by $\mathbb{Z}$; the nonnegative integers by $\mathbb{Z}_+$; the real rationals by $\mathbb{Q}$ and the natural numbers $1, 2, \ldots$ by $\mathbb{N}$. As usual, $\mathbb{R}$ and $\mathbb{C}$ denote the real line and the complex plane, respectively, but also the real and complex number fields. The symbol $\mathbb{R}_+$ means the half-line $[0, \infty)$. If $z \in \mathbb{C}$ then $z = \Re(z) + \Im(z)\sqrt{-1}$ where $\Re(z)$ and $\Im(z)$ are real. Since both fields $\mathbb{R}$ and $\mathbb{C}$ are needed in definitions, let the symbol $\mathbb{F}$ indicate either field. $\mathbb{F}^n$ will denote the n-dimensional linear space of column vectors $x = \mathrm{col}(x_1, \ldots, x_n)$ whose components are $x_i \in \mathbb{F}$. Linear and rational symbolic calculations use the field of real rationals $\mathbb{Q}$ or an algebraic extension of $\mathbb{Q}$ such as the complex rationals $\mathbb{Q}(\sqrt{-1})$.

The Kronecker delta symbol $\delta_{i,j}$ is defined by $\delta_{i,j} := 1$ if $i = j$, $\delta_{i,j} := 0$ if $i \neq j$. A standard basis for $\mathbb{F}^n$ is

$$\{\delta_p := \mathrm{col}(\delta_{1,p}, \ldots, \delta_{n,p}) |\, p \in 1\,..\,n\}, \text{ or explicitly}$$
$$\delta_1 := \mathrm{col}(1, 0, \ldots, 0), \ldots, \delta_n := \mathrm{col}(0, 0, \ldots, 1). \tag{A.1}$$

Roman capital letters $A, B, \ldots, Z$ are used for linear mappings $\mathbb{F}^n \to \mathbb{F}^n$ and their representations as square matrices in the linear space $\mathbb{F}^{n\times n}$; *linear* means $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$ for all $\alpha, \beta \in \mathbb{F}$, $x, y \in \mathbb{F}^n$, $A \in \mathbb{F}^{n\times n}$. The set $\{x \in \mathbb{F}^n |\, Ax = 0\}$ is called the nullspace of A; the range of A is $\{Ax |\, x \in \mathbb{F}^n\}$. The entries of A are by convention $\{a_{i,j} |\, i, j \in 1, \ldots, n\}$, so one may write A as $\{a_{i,j}\}$ except that the identity operator on $\mathbb{F}^n$ is represented by $I_n := \{\delta_{i,j}\}$, or I if n is understood.[1] An $i \times j$ matrix of zeros is denoted by $\mathbf{0}_{i,j}$.

Much of the matrix algebra needed for control theory is given in Sontag [249, Appendix A] and in Horn and Johnson [131, 132].

[1] The symbol I means $\sqrt{-1}$ to Mathematica; beware.

D.L. Elliott, *Bilinear Control Systems,* Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_9,

A.1.1 Some Associative Algebras

An $\mathbb{F}$-linear space $\mathfrak{A}$ equipped with a multiplication $* : \mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$ is called an associative algebra over $\mathbb{F}$ if $*$ satisfies the associative law and distributes over addition and scalar multiplication. We use two important examples:

(i) The linear space of $n \times n$ matrices $\mathbb{F}^{n\times n}$, with $*$ the usual matrix multiplication $A, B \to AB$, $\mathbb{F}^{n\times n}$ is an associative matrix algebra with identity I_n. See Section A.2.
(ii) The linear space $\mathbb{F}[x_1, \cdots, x_n]$ of polynomials in n indeterminates with coefficients in $\mathbb{F}$ can be given the structure of an associative algebra $\mathfrak{A}$ with $*$ the usual multiplication of polynomials. For symbolic calculation, a total order must be imposed on the indeterminates and their powers,[2] so that there is a one-to-one mapping of polynomials to vectors; see Section C.1.

A.1.2 Operations on Matrices

As usual, the trace of A is $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{i,i}$; its determinant is denoted by $\det(A)$ or $|A|$ and the inverse of A by A^{-1}; of course $A^0 = I$. The symbol $\sim$, which generally is used for equivalence relations, will indicate similarity for matrices: $\tilde{A} \sim A$ if $\tilde{A} = P^{-1}AP$ for some invertible matrix P.

A dth-order minor, or d–minor, of a matrix A is the determinant of a square submatrix of A obtained by choosing the elements in the intersection of d rows and d columns; a principal minor of whatever order has its top left and bottom right entries on the diagonal of A.

The transpose of $A = \{a_{i,j}\}$ is $A^\top = \{a_{j,i}\}$; in consistency with that, $x^\top$ is a row vector. If $Q \in \mathbb{R}^{n\times n}$ is real and $Q^\top = Q$, then Q is called symmetric and we write $Q \in \mathrm{Symm}(n)$. Given any matrix $A \in \mathbb{R}^{n\times n}$, $A = B + C$ where $B \in \mathrm{Symm}(n)$ and $C + C^\top = 0$; C is called skew-symmetric or antisymmetric.

If $z = x + iy$, $x = \Re z$, $y = \Im z; \bar{z} = x - iy$. As usual $z^* := \bar{z}^\top$. The usual identification of $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ is given by

$$\begin{bmatrix} x_1 + iy_1 \\ \vdots \\ x_n + iy_n \end{bmatrix} \leftrightarrow \mathrm{col}(x_1, \ldots, x_n,\ y_1, \ldots, y_n). \tag{A.2}$$

The complex-conjugate transpose of A is $A^* := \bar{A}^\top$, so if A is real $A^* = A^\top$. We say $Q \in \mathbb{C}^{n\times n}$ is a Hermitian matrix, and write $Q \in \mathrm{Herm}(n)$, if $Q^* = Q$. If $Q^* + Q = 0$ one says Q is skew-Hermitian. If $Q \in \mathrm{Herm}(n)$ and $Qz = \lambda z$, then $\lambda \in \mathbb{R}$. Matrix A is called unitary if $A^*A = I$ and orthogonal if $A^\top A = I$.

---

[2] For the usual total orders, including lexicographic order, and where to use them see Cox et al. [67].

A.1.3 Norms

The Euclidean vector norm for $x \in \mathbb{F}^n$ and its induced operator norm on $\mathbb{F}^{n\times n}$ are

$$\|x\| = \left(\sum_i^n |x_i|^2\right)^{1/2}, \ \|A\| = \sup_{\|x\|=1} \|Ax\|.$$

A sphere of radius $\epsilon > 0$ is denoted by $S_\epsilon = \{x| \, \|x\| = \epsilon\}$. It is straightforward to verify the following inequalities ($\alpha \in \mathbb{F}$):

$$\|A + B\| \leq \|A\| + \|B\|, \ \|AB\| \leq \|A\| \, \|B\|, \ \|\alpha A\| \leq |\alpha| \, \|A\|. \tag{A.3}$$

The Frobenius norm is $\|A\|_2 = (\operatorname{tr} AA^*)^{1/2}$; over $\mathbb{R}$, $\|A\|_2 = (\operatorname{tr} AA^\top)^{1/2}$. It satisfies the same inequalities and induces the same topology on matrices as $\| \cdot \|$ because $\|A\| \leq \|A\|_2 \leq \sqrt{n}\|A\|$. The related inner product for complex matrices is $\langle A, B\rangle = \operatorname{tr} AA^*$; for real matrices it is $\langle A, B\rangle = \operatorname{tr} AA^\top$.

A.1.4 Eigenvalues

Given a polynomial $p(z) = \sum_1^N c_i z^i$, one can define the corresponding matrix polynomial on $\mathbb{F}^{n\times n}$ by $p(A) = \sum_1^N c_i A^i$. The characteristic polynomial of a square matrix A is denoted by

$$\begin{aligned} p_A(\lambda) :=& \det(\lambda I - A) = \lambda^n + \cdots + \det A \\ =& (\lambda - \alpha_1)(\lambda - \alpha_2) \cdots (\lambda - \alpha_n); \end{aligned}$$

the set of its n roots is called the spectrum of A and is denoted by $\operatorname{spec}(A) := \{\alpha_1, \ldots, \alpha_n\}$. The α_i are called eigenvalues of A.

A.2 Associative Matrix Algebras

A linear subspace $\mathfrak{A}^n \subset \mathbb{F}^{n\times n}$ that is also closed under the matrix product is called an associative matrix algebra. One example is $\mathbb{F}^{n\times n}$ itself. Two associative matrix algebras $\mathfrak{A}^n, \mathfrak{B}^n$ are called conjugate if there exists a one-to-one correspondence $\mathfrak{A}^n \leftrightarrow \mathfrak{B}^n$ and a nonsingular $T \in \mathbb{F}^{n\times n}$ such that for a fixed T corresponding matrices A, B are similar, $TAT^{-1} = B$. Given a matrix basis $B_1, \ldots, B_d$ for $\mathfrak{A}^n$, the algebra has a multiplication table, preserved under conjugacy,

$$B_i B_j = \sum_{k=1}^d \beta_{i,j}^k B_k, \ i, j, k \in 1, \ldots, d.$$

A matrix $A \in \mathbb{F}^{n\times n}$ is called nilpotent if $A^n = 0$ for some $n \in \mathbb{N}$. A is called unipotent if $A - I$ is nilpotent. A unipotent matrix is invertible and its inverse is unipotent. An associative algebra $\mathfrak{A}$ is called nilpotent (unipotent) if all of its elements are nilpotent (unipotent).

Given a list of matrices $\{A_1, \dots, A_k\}$, linearly independent over $\mathbb{F}$, define $\{A_1, \dots, A_k\}_{\mathcal{A}}$ to be the smallest associative subalgebra of $\mathbb{F}^{n\times n}$ containing them; the A_i are called generators of the subalgebra. If $A_iA_j = A_jA_i$ one says that the matrices A_i and A_j commute, and all of its elements commute the subalgebra is called Abelian or commutative. Examples:

- The set of all $n \times n$ diagonal matrices $\mathfrak{D}(n, \mathbb{F})$ is an associative subalgebra of $\mathbb{F}^{n\times n}$ with identity I; a basis of $\mathfrak{D}(n, \mathbb{F})$ is

$$B_1 = \{\delta_{i,1}\delta_{i,j}\}, \dots, B_n = \{\delta_{i,n}\delta_{i,j}\};$$

  $\mathfrak{D}(n, \mathbb{F})$ is an Abelian algebra whose multiplication table is $B_iB_j = \delta_{i,j}B_i$.
- Upper triangular matrices are those whose subdiagonal entries are zero. The linear space of strictly upper-triangular matrices

$$\mathfrak{n}(n, \mathbb{F}) := \{A \in \mathbb{F}^{n\times n} | \, a_{i,j} = 0, \ i \le j\}$$

  under matrix multiplication is an associative algebra with no identity element; all elements are nilpotent.

A.2.1 Cayley–Hamilton

The Cayley–Hamilton Theorem over $\mathbb{F}$ is proved in most linear algebra textbooks. The following general version is also well known, see Brown [41].

Theorem A.1 (Cayley–Hamilton). *If* $\mathbf{R}$ *is a commutative ring with identity* 1, $A \in R^{n\times n}$, *and* $\boldsymbol{p}_A(t) := \det(tI_n - A)$ *then* $\boldsymbol{p}_A(A) = 0$.

Discussion Postulate the theorem for $\mathbf{R} = \mathbb{C}$. The entries $p_{i,j}$ of the matrix $P := \boldsymbol{p}_A(A)$ are formal polynomials in the indeterminates $a_{i,j}$, with integer coefficients. Substituting arbitrary complex numbers for the $a_{i,j}$ in $\boldsymbol{p}_A(A)$, for every choice $p_{i,j} \equiv 0$; so $p_{i,j}$ is the zero polynomial, therefore $P = 0$.[3] As an example, let $1, \alpha, \beta, \gamma, \delta$ be the generators of $\mathbf{R}$;

$$A := \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}; \ \boldsymbol{p}_A(A) = A^2 - (\alpha + \delta)A + (\alpha\delta - \beta\gamma)I \equiv \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \qquad \triangle$$

[3] This proof by substitution seems to be folklore of algebraists.

A.2.2 Minimum Polynomial

For a given $A \in \mathbb{F}^{n\times n}$, the associative subalgebra $\{I, A\}_{\mathcal{A}}$ consists of all the matrix polynomials in A. As a corollary of the Cayley–Hamilton Theorem there exists a monic polynomial P_A with minimal degree κ such that $P_A(A) = 0$, known as the minimum polynomial of A. Let $\kappa = \deg P_A$. Since any element of $\{I, A\}_{\mathcal{A}}$ can be written as a polynomial $f(A)$ of degree at most $\kappa - 1$, $\dim\{I, A\}_{\mathcal{A}} = \kappa$.

Remark A.1. The associative algebra $\{1, s\}_{\mathfrak{A}}$ generated by one indeterminate s can be identified with the algebra of formal power series in one variable; if it is mapped into $\mathbb{F}^{n\times n}$ with a mapping that assigns $1 \to I,\ s \to A$ and preserves addition and multiplication (an algebra homomorphism) the image is $\{I, A\}_{\mathcal{A}}$.
△

A.2.3 Triangularization

Theorem A.2 (Schur's Theorem). *For any $A \in \mathbb{F}^{n\times n}$ there exists $T \in \mathbb{C}^{n\times n}$ such that $\tilde{A} := T^{-1}AT$ is in upper triangular form with its eigenvalues λ_i along the diagonal. Matrix T can be chosen to be unitary: $TT^* = I$.*[4]

If A and its eigenvalues are real, so are T and $\tilde{A}$; $\tilde{A}$ can be chosen to be the block-diagonal Jordan canonical form, with s blocks of size $k_i \times k_i$ ($\sum_1^s k_i = n$) with superdiagonal 1's. Having Jordan blocks of size $k_i > 1$ is called the resonance case.

$$\tilde{A} = \begin{bmatrix} J_1 & 0 & \cdots \\ & \ddots & \\ 0 & \cdots & J_s \end{bmatrix}, \quad J_k = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots \\ & \ddots & \ddots & \\ 0 & \cdots & \lambda_i & 1 \\ 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

One of the uses of the Jordan canonical form is in obtaining the widest possible definition of a function of a matrix, especially the logarithm.

Theorem A.3 (Horn and Johnson [132], Th. 6.4.15). *Given a real matrix A, there exists real F such that* $\exp(F) = A$ *if and only if A is nonsingular and has an even number of Jordan blocks of each size for every negative eigenvalue in* $\operatorname{spec}(A)$.
△

The most common real canonical form A^R of a matrix A has its real eigenvalues along the diagonal, while a complex eigenvalue $\lambda = \alpha + \beta\sqrt{-1}$ is represented by a real block $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ on the diagonal as in this example where complex λ and $\bar{\lambda}$ are double roots of A:

[4] See Horn and Johnson [131, Th. 2.3.11] for the proof.

$$A = \begin{bmatrix} \lambda & 0 & 1 & 0 \\ 0 & \bar{\lambda} & 0 & 1 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \bar{\lambda} \end{bmatrix}, \quad A^R = \begin{bmatrix} \alpha & \beta & 1 & 0 \\ -\beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{bmatrix}.$$

A.2.4 Irreducible Families

A family of matrices $\mathcal{F} \subset \mathbb{F}^{n\times n}$ is called irreducible if there exists no nontrivial linear subspace of $\mathbb{F}^n$ invariant under all of the mappings in $\mathcal{F}$. Such a family may be, for instance, a matrix algebra or group. The centralizer in $\mathbb{F}^{n\times n}$ of a family of matrices $\mathcal{F}$ is the set $\mathfrak{z}$ of matrices X such that $AX = XA$ for every $A \in \mathcal{F}$; it contains, naturally, the identity.

Lemma A.1 (Schur's Lemma).[5] *Suppose $\mathcal{F}$ is a family of operators irreducible on $\mathbb{C}^n$ and $\mathfrak{z}$ is its centralizer. Then $\mathfrak{z} = \mathbb{C}I_n$.*

Proof. For any matrix $Z \in \mathfrak{z}$ and for every $A \in \mathcal{F}$, $[Z, A] = 0$. The matrix Z has an eigenvalue $\lambda \in \mathbb{C}$ and corresponding eigenspace V_λ, which is Z-invariant. For $x \in V_\lambda$ we have $AZx = \lambda Ax$, so $Z(Ax) = \lambda Ax$ hence $Ax \in V_\lambda$ for all $A \in \mathcal{F}$. That is, V_λ is $\mathcal{F}$-invariant but by irreducibility it has no nontrivial invariant subspace, so $V_\lambda = \mathbb{C}^n$ and $\mathfrak{z} = \mathbb{C}I_n$. □

A.3 Kronecker Products

The facts about Kronecker products, sums, and powers given here may be found in Horn and Johnson [132, Ch. 4] and Rugh [228, Ch. 3]; also see Bellman [25, Appendix D]. The easy but notationally expansive proofs are omitted.

Given $m\times n$ matrix $A = \{a_{i,j}\}$ and $p\times q$ matrix $B = \{b_{i,j}\}$, then their Kronecker product is the matrix $A \otimes B := \{a_{i,j}B\} \in \mathbb{F}^{mp\times nq}$; in detail, with $\nu := n^2$,

$$A \otimes B = \begin{bmatrix} a_{1,1}B & a_{1,2}B & \dots & a_{1,n}B \\ \dots & \dots & \dots & \dots \\ a_{m,1}B & a_{m,2}B & \dots & a_{m,n}B \end{bmatrix}. \tag{A.4}$$

$$\text{If } x \in \mathbb{F}^n, y \in \mathbb{F}^p,\ x \otimes y := \mathrm{col}(x_1 y, x_2 y, \dots, x_n y) \in \mathbb{F}^{np}. \tag{A.5}$$

[5] The statement and proof of Lemma A.1 are true for $\mathbb{R}^n$ if n is odd because in that case any Z must have a real eigenvalue.

A.3.1 Properties

Using partitioned multiplication Horn and Johnson [132, Lemma 4.2.10] shows that for compatible pairs A, C and B, D

$$(AC) \otimes (BD) = (A \otimes B)(C \otimes D). \tag{A.6}$$

Generally $A \otimes B \neq B \otimes A$, but from (A.6) $(I \otimes A)(B \otimes I) = B \otimes A = (B \otimes I)(I \otimes A)$.

Now specialize to A, B in $\mathbb{F}^{n\times n}$. The Kronecker product $A \otimes B \in \mathbb{F}^{\nu\times\nu}$ has n^4 elements and is bilinear, associative, distributive over addition, and noncommutative. As a consequence of (A.6) if A^{-1} and B^{-1} exist then $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$. If A and B are upper (lower) triangular, so is $A \otimes B$. Given

$$\begin{aligned}&A, B \in \mathbb{F}^{n\times n},\ \mathrm{spec}(A) = \{\alpha_1, \ldots, \alpha_n\},\text{ and } \mathrm{spec}(B) = \{\beta_1, \ldots, \beta_n\},\text{ then}\\ &\mathrm{spec}(A \otimes B) = \mathrm{spec}(A)\,\mathrm{spec}(B) = \{\alpha_1\beta_1, \alpha_1\beta_2, \ldots, \alpha_{n-1}\beta_n, \alpha_n\beta_n\}.\end{aligned} \tag{A.7}$$

A.3.2 Matrices as Vectors

Let $\nu := n^2$. Any matrix $A = \{a_{i,j}\} \in \mathbb{F}^{n\times n}$ can be $\mathbb{F}$-linearly mapped one-to-one to a vector $A^\flat \in \mathbb{F}^\nu$ by stacking[6] its columns in numerical order:

$$A^\flat := \mathrm{col}(a_{1,1}, a_{2,1}, \ldots, a_{n,1}, a_{1,2}, a_{2,2}, \ldots, a_{n,n}). \tag{A.8}$$

Pronounce $A^\flat$ as "A flat." The inverse operation to $(\cdot)^\flat$ is $\sharp(\cdot): \mathbb{R}^\nu \to \mathbb{R}^{n\times n}$ ($\sharp$ is pronounced *sharp*, of course), which partitions a ν-vector into an $n \times n$ matrix; $\sharp(A^\flat) = A$. The Frobenius norm $\|A\|_2$ (Section A.1.3) is given by $\|A\|_2^2 = (A^\flat)^\top A^\flat$.

The key fact about flattened matrices $A^\flat$ and the Kronecker product is that for compatible matrices A, Y, B (square matrices, in this book)

$$(AYB)^\flat = (B^\top \otimes A)Y^\flat. \tag{A.9}$$

A.3.3 Sylvester Operators

Chapters 1 and 2 make use of special cases of the Sylvester operator $\mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n} : X \to AX + XB$ such as the Lyapunov operator $\mathrm{Ly}_A(X) := A^\top X + XA$ and

---

[6] Horn and Johnson [132] use the notation $\mathbf{vec}(A)$ for $A^\flat$. Mathematicastores a matrix A row-wise as a list of lists in lexicographic order $\{\{a_{1,1}, a_{1,2}, \ldots,\}, \{a_{2,1},\}, \ldots\}$ which `Flatten` flattens to a single list $\{a_{1,1}, a_{1,2}, \ldots, a_{2,1}, a_{2,2} \ldots\}$, so if using Mathematica to do Kronecker calculus compatible with [132], to get $a = A^\flat$ use
`a = Flatten[Transpose[A]].`

the adjoint operator $\mathrm{ad}_A(X) := AX - XA$. From (A.9)

$$(AX + XB)^\flat = (I \otimes A + B^\top \otimes I)X^\flat, \tag{A.10}$$

$$(\mathrm{Ly}_A(X))^\flat = (I \otimes A^\top + A^\top \otimes I)X^\flat, \tag{A.11}$$

$$(\mathrm{ad}_A(X))^\flat = (I \otimes A - A^\top \otimes I)X^\flat. \tag{A.12}$$

The operator in (A.10) is called the Kronecker sum of A and B, and will be denoted[7] here as $A \boxplus B := I \otimes A + B^\top \otimes I$. Horn and Johnson [132, Ch. 4] is a good resource for Kronecker products and sums, also see Bellman [25].

Proposition A.1. *[132, Th. 4.4.6]. Sylvester's equation $AX + XB = C$ has a unique solution X for each $C \in \mathbb{F}^{n\times n}$ if and only if* $\mathrm{spec}(A) \cap \mathrm{spec}(-B) = \emptyset$.

Proof. The eigenvalues of $A \boxplus B$ are $\mu_{i,j} = \alpha_i + \beta_j$ for each pair $\alpha_i \in \mathrm{spec}(A)$, $\beta_j \in \mathrm{spec}(B)$. □

A.3.4 Kronecker Powers

The Kronecker powers of $X \in \mathbb{F}^{m\times n}$ are

$$X^{\otimes 1} := X,\ X^{\otimes 2} := X \otimes X, \dots, X^{\otimes k} := X \otimes X^{\otimes k-1} \in \mathbb{F}^{m^k \times n^k}. \tag{A.13}$$

Example: for $x \in \mathbb{F}^n \simeq \mathbb{F}^{n\times 1}$, $x^{\otimes 2} = \mathrm{col}(x_1^2, \dots, x_1 x_n, x_2 x_1, \dots, x^2 x^n, \dots, x_n^2)$.

A special notation will be used in Proposition A.2 and in Section 5.7 to denote this $n^{j+1} \times n^{j+1}$ matrix:

$$\sigma(A, j) := \left(A \otimes I^{\otimes j} + I \otimes A \otimes I^j + \cdots + I^{\otimes j} \otimes A\right). \tag{A.14}$$

Proposition A.2. *Given $\dot{x} = Ax$ and $z = x^{\otimes k}$ then $\dot{z} = \sigma(A, k-1)z$.*

Proof (Sketch). The product rule for differentiation follows from the multilinearity of the Kronecker product; everything else follows from the repeated application of (A.6). Thus we have

$$\begin{aligned}
\frac{d}{dt}x^{\otimes 2} &= Ax \otimes Ix + Ix \otimes Ax = (A \otimes I) \otimes x^{\otimes 2} + (I \otimes A) \otimes x^{\otimes 2} = \sigma(A,1)x^{\otimes 2};\\
\frac{d}{dt}x^{\otimes 3} &= (Ax \otimes Ix) \otimes x + x \otimes (Ax \otimes Ix) + x \otimes (Ix \otimes Ax)\\
&= (A \otimes I) \otimes x^{\otimes 2} \otimes Ix + Ix \otimes (A \otimes I)x^{\otimes 2} + Ix \otimes (I \otimes A)x^{\otimes 2}\\
&= (A \otimes I^{\otimes 2})x^{\otimes 3} + (I^{\otimes 2} \otimes A)x^{\otimes 3} + I^{\otimes 3}x^{\otimes 3} = \sigma(A,2)x^{\otimes 3}.
\end{aligned}$$

[7] The usual notation for $A \boxplus B$ is $A \oplus B$, which might be confusing in view of Section B.1.4.

To complete a proof by induction requires rewriting the general term $x^{\otimes i-1} \otimes Ax \otimes x^{\otimes k-i}$ using (A.6) and identities already obtained. $\square$

A.4 Invariants of Matrix Pairs

The conjugacy

$$(B_1, \ldots, B_m) \to (P^{-1}B_1P, \ldots, P^{-1}B_mP)$$

is an action of $\mathrm{GL}(n, \mathbb{F})$ on $(\mathbb{F}^{n\times n})^m$. The general problem of classifying m-tuples of matrices up to conjugacy under $\mathrm{GL}(n, \mathbb{F})$ has been studied for many years and from many standpoints; see Chapter 3. Procesi [221] found that the invariants of $B_1, \ldots, B_m$ under conjugacy are all polynomials in traces of matrix products; these can be listed as

$$\{\mathrm{tr}(B_{i_1} \ldots B_{i_k}) | \, k \leq 2^n - 1\}. \tag{A.15}$$

The strongest results are for $m = 2$; Friedland [100] provided a complete classification for matrix pairs $\{A, B\}$. The product-traces in (A.15) appear because they are elementary symmetric polynomials, sometimes called power-sums, of the roots of the characteristic polynomial $\boldsymbol{p}_{A,B}(s,t) := \det(sI - (A + tB))$; using Newton's identities[8] its coefficients can be expressed as polynomials in the product-traces.

A.4.1 The Second Order Case

As an example, in Friedland [100] the case $n = 2$ is worked out in detail. The identity $2\det(X) = \mathrm{tr}(X)^2 - \mathrm{tr}(X^2)$ is useful in simplifying the characteristic polynomial in this case:

$$\begin{aligned} \boldsymbol{p}_{A,B}(s,t) &:= s^2 - s(\mathrm{tr}(A) + t\,\mathrm{tr}(B)) + \det(A + tB) \\ &= s^2 - s\,\mathrm{tr}(A) + \frac{1}{2}(\mathrm{tr}(A)^2 - \mathrm{tr}(A^2)) - st\,\mathrm{tr}(B) \qquad \text{(A.16)} \\ &\quad + t(\mathrm{tr}(A)\,\mathrm{tr}(B) - \mathrm{tr}(AB)) + \frac{t^2}{2}(\mathrm{tr}(B)^2 - \mathrm{tr}(B^2)) \\ &= s^2 + a_1 s + a_2 st + b_0 + b_1 t + b_2 t^2. \qquad \text{(A.17)} \end{aligned}$$

Paraphrasing [100], the coefficients of (A.16) (or (A.17)) are a complete set of invariants for the orbits of matrix pairs *unless* A, B satisfy this hypothesis: (H) $\boldsymbol{p}_{A,B}$ can be factored over $\mathbb{C}[s,t]$ into two linear terms.

[8] For symmetric polynomials and Newton identities see Cox et al. [67, Ch. 7].

If (H) then additional classifying functions, rational in the product-traces, are necessary (but will not be listed here); and A, B are simultaneously similar to upper triangular matrices ($\{A, B\}_{\mathcal{L}}$ is solvable).

Friedland [100] shows that (H) holds on an affine variety U in $\mathbb{C}^{2\times 2} \times \mathbb{C}^{2\times 2}$ given by $\Delta = 0$ where using product traces and (A.17)

$$\Delta := (2\,\mathrm{tr}(A^2) - \mathrm{tr}(A)^2)(2\,\mathrm{tr}(B^2) - \mathrm{tr}(B)^2) - (2\,\mathrm{tr}(AB) - \mathrm{tr}(A)\,\mathrm{tr}(B))^2 \tag{A.18}$$

$$= 16(b_0 a_2^2 - a_1 a_2 b_1 + b_1^2 - 4 b_0 b_2 + a_1^2 b_2). \tag{A.19}$$

Exercise A.1. Verify $\Delta = \mathrm{Discriminant}(\mathrm{Discriminant}(\boldsymbol{p}_{A,B}, s), t)$ where in Mathematica, for any n

```
 Discriminant[p_,x_]:=Module[{m,dg,den,temp},
m=Exponent[p,x]; dg=(-1)^(m*(m-1)/2);
den=Coefficient[p,x,m];
Cancel[dg*Resultant[p, D[p,x],x]/den] ]
```

$\triangle$

Appendix B
Lie Algebras and Groups

This chapter, a supplement to Chapter 2, is a collection of standard definitions and basic facts drawn from Boothby [31], Conlon [64], Jacobson [143], Jurdjevic [147], and especially Varadarajan [282].

B.1 Lie Algebras

The field $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. If an $\mathbb{F}$-linear space $\mathfrak{g}$ of any finite or infinite dimension ℓ is equipped with a binary operation $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} : (e, f) \mapsto [e, f]$ (called the Lie bracket) satisfying Jacobi's axioms

$$[e, \alpha f + \beta g] = \alpha[e, f] + \beta[e, g] \quad \text{for all } \alpha, \beta \in \mathbb{F}, \tag{Jac.1}$$

$$[e, f] + [f, e] = 0, \tag{Jac.2}$$

$$[e, [f, g]] + [f, [g, e]] + [g, [e, f]] = 0, \tag{Jac.3}$$

then $\mathfrak{g}$ will be called a Lie algebra over $\mathbb{F}$. In this book, the dimension ℓ of $\mathfrak{g}$ is assumed to be finite unless an exception is noted.

B.1.1 Examples

Any ℓ-dimensional linear space L can trivially be given the structure of an Abelian Lie algebra by defining $[e, f] = 0$ for all $e, f \in L$.

The space of square matrices $\mathbb{F}^{n\times n}$ becomes a Lie algebra if the usual matrix multiplication is used to define $[A, B] := AB - BA$.

Partial differential operators and multiplier functions that act on differentiable functions ψ are important in quantum physics. They sometimes generate low-dimensional Lie algebras; see Wei and Norman [284]. One such is the Heisenberg algebra $\mathfrak{g}_Q = \text{span}\{1, p, q\}$ with the relations $[1, p] = 0 = [1, q]$,

D.L. Elliott, *Bilinear Control Systems,* Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_10,

$[p, q] = 1$. One of its many representations uses the position multiplier x and momentum operator $p = \partial/\partial x$, with $\partial(x\psi)/\partial x - x\partial\psi/\partial x = \psi$.

B.1.2 Subalgebras; Generators

Given a Lie algebra $\mathfrak{g}$, a subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is a linear subspace of $\mathfrak{g}$ that contains the Lie bracket of every pair of elements, that is, $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$. If also $[\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h}$, the subalgebra $\mathfrak{h}$ is called an ideal of $\mathfrak{g}$. Given a list of elements (generators) $\{e, f, \dots\}$ in a Lie algebra $\mathfrak{g}$ define $\{e, f, \dots\}_{\mathcal{L}} \subset \mathfrak{g}$ to be the smallest Lie subalgebra that contains the generators and is closed under the bracket and linear combination. Such generated Lie subalgebras are basic to bilinear control system theory.

B.1.3 Isomorphisms

An isomorphism of Lie algebras $\mathfrak{g}$, $\mathfrak{h}$ over $\mathbb{F}$ is a one-to-one map

$$\pi: \ \mathfrak{g} \to \mathfrak{h} \text{ such that for all } e, f \in \mathfrak{g}$$
$$\pi(\alpha e + \beta f) = \alpha\pi(e) + \beta\pi(f), \ \alpha, \beta \in \mathbb{F}, \tag{i}$$
$$\pi([e, f]_1) = [\pi(e), \pi(f)]_2, \tag{ii}$$

where $[\cdot,\cdot]_1$, $[\cdot,\cdot]_2$ are the Lie bracket operations in $\mathfrak{g}_1$, $\mathfrak{g}_2$, respectively. If π is not one-to-one but is onto and satisfies (i) and (ii), it is called a Lie algebra homomorphism. One type of isomorphism for matrix Lie algebras is conjugacy: $P^{-1}\mathfrak{g}_1 P = \mathfrak{g}_2$.

B.1.4 Direct Sums

Let the Lie algebras $\mathfrak{g}$, $\mathfrak{h}$ be complementary as vector subspaces of $\mathfrak{gl}(n, \mathbb{R})$ with the property that if $X \in \mathfrak{g}$, $Y \in \mathfrak{h}$ then $[X, Y] = 0$. Then

$$\mathfrak{g} \oplus \mathfrak{h} := \{\alpha X + \beta Y | \, X \in \mathfrak{g}, \ Y \in \mathfrak{h}; \alpha, \beta \in \mathbb{R}\}$$

is a Lie algebra called the (internal) direct sum of $\mathfrak{g} \oplus \mathfrak{h}$. See Appendix D for examples; see Definition B.15 for the corresponding notion for Lie groups, the direct product.

B.1.5 Representations and Matrix Lie Algebras

A representation (faithful representation) $(\varrho, \mathbb{F}^n)$ of a Lie algebra $\mathfrak{g}$ on $\mathbb{F}^n$ is a Lie algebra homomorphism (isomorphism) into (onto) a Lie subalgebra $\varrho(\mathfrak{g}) \subset \mathfrak{gl}(n, \mathbb{F})$; such a subalgebra is called a matrix Lie algebra. See Theorem B.2 (Ado's Theorem).

*B.1.6 Free Lie Algebras**

The so-called free Lie algebra $\mathfrak{L}_r$ on $r \geq 2$ generators is a *Lie ring* (not a Lie algebra)[1] over $\mathbb{Z}$ defined by a list $\{\lambda_1, \lambda_2, \dots, \lambda_r\}$ of indeterminates (generators) and two operations. These are addition (yielding multiplication by integers) and a formal Lie bracket satisfying Jac.1—Jac.3. Note that $\mathfrak{L}_r$ contains all the repeated Lie products of its generators, such as

$$[\lambda_1, \lambda_2],\ [\lambda_3, [\lambda_1, \lambda_2]],\ [[\lambda_1, \lambda_2], [\lambda_3, \lambda_1]], \dots,$$

which are called Lie monomials, and their formal $\mathbb{Z}$-linear combinations are called Lie polynomials. There are only a finite number of monomials of each *degree* ($\deg(\lambda_i) = 1, \deg([\lambda_i, \lambda_j]) = 2, \dots$). The set $\mathcal{Q}_r$ of standard monomials q on r generators can be defined recursively. First we define a linear order:

$q_1 \prec q_2$ if $\deg(q_1) < \deg(q_2)$;

if $\deg(q_1) = \deg(q_2)$ then q_1 and q_2 are ordered lexicographically.

For example, $\lambda_1 \prec \lambda_2$, $\lambda_2 \prec [\lambda_1, \lambda_2] \prec [\lambda_2, \lambda_1]$. We define $\mathcal{Q}_r$ by (i) $\lambda_i \in \mathcal{Q}_r$ and (ii) if $q_1, q_2 \in \mathcal{Q}_r$ and $q_1 \prec q_2$, append $[q_1, q_2]$ to $\mathcal{Q}_r$; if $q_2 \prec q_1$, append $[q_2, q_1]$.

Theorem B.1 (M. Hall). *For each d the standard monomials of degree d are a basis $\mathbf{B}^{(d)}$ for the linear space L_r^d of homogeneous Lie polynomials of degree d in r generators. The union over all $d \geq 1$ of the $\mathbf{B}^{(d)}$ is a countably infinite basis $\mathcal{PH}(r)$, called a Philip Hall basis of $\mathfrak{L}_r$.*

This (with a few differences in the definitions) is stated and proved in a paper of Marshall Hall, Jr. [116, Th. 3.1] (*cf.* [78] or [143, V.4]).

The dimension $d(n, r)$ of L_r^n is given by the Witt dimension formula

[1] There is no linear space structure for the generators; there are no relations among the generators, so $\mathfrak{L}_r$ is "free." See Serre [240] and for computational purposes see Reutenauer [223].

$$d(n,r) = \frac{1}{n}\sum_{d|n}\mu(d)r^{n/d} \text{ where } \mu \text{ is the Möbius function.}$$

$$\mu(d) := \begin{cases} 1, & d = 1, \\ 0, & \text{if } d \text{ has a square factor,} \\ (-1)^k, & \text{if } d = p_1 \dots p_k \text{ where the } p_i \text{ are prime numbers.} \end{cases}$$

For $r = 2$

n :	1	2	3	4	5	6
$d(n,2)$:	2	1	2	3	6	9

The standard monomials can be arranged in a tree as in Table B.1 (its construction takes into account identities that arise from the Jacobi axioms). Note that after the first few levels of the tree, many of the standard monomials are not *left normalized* (defined in Section 2.2.4).

		α	β	$\mathbf{B}^{(1)}$
		$[\alpha,\beta]$		$\mathbf{B}^{(2)}$
	$[\alpha,[\alpha,\beta]]$		$[\beta,[\alpha,\beta]]$	$\mathbf{B}^{(3)}$
$[\alpha,[\alpha,[\alpha,\beta]]]$	$[\beta,[\alpha,[\alpha,\beta]]]$		$[\beta,[\beta,[\alpha,\beta]]]$	$\mathbf{B}^{(4)}$
$[\alpha,[\alpha,[\alpha,[\alpha,\beta]]]]$	$[\beta,[\alpha,[\alpha,[\alpha,\beta]]]]$	$[\beta,[\beta,[\alpha,[\alpha,\beta]]]]$	$[\beta,[\beta,[\beta,[\alpha,\beta]]]]$	
		$[[\alpha,\beta],[\alpha,[\alpha,\beta]]]$	$[[\alpha,\beta],[\beta,[\alpha,\beta]]]$	$\mathbf{B}^{(5)}$

Table B.1. Levels $\mathbf{B}^{(1)}$–$\mathbf{B}^{(5)}$ of a Philip Hall basis of $\mathfrak{L}_2$ with generators α, β.

Torres-Torriti and Michalska [278] provides algorithms that generate $\mathfrak{L}_m$. Given any Lie algebra $\mathfrak{g}$ with m generators (for instance, nonlinear vector fields) and a Lie homomorphism $h : \mathfrak{L}_m \to \mathfrak{g}$ that takes the generators of $\mathfrak{L}_m$ to the generators of $\mathfrak{g}$, the image of h is a set of Lie monomials whose real span is $\mathfrak{g}$.

Caveat: given matrices $B_1, \dots, B_m$, the free Lie algebra $\mathfrak{L}_m$ and each level $\mathbf{B}(k)$ of the Hall basis $\mathcal{PH}(m)$ can be mapped homomorphically by h into $\{B_1, \dots, B_m\}_{\mathcal{L}}$, but for $k > 2$ the image $h(\mathbf{B}(k))$ will likely contain linearly dependent matrices. Free Lie algebras and Philip Hall bases are interesting and are encountered often enough to warrant including this section. There is a general impression abroad that the generation of Lie algebra bases should take advantage of them. The efficient matrix algorithm of Isidori [140] that is implemented in Table 2.2 makes no use of Philip Hall bases.

B.1.7 Structure Constants

If a basis $\{e^1, e^2, \ldots, e^\ell\}$ is specified for Lie algebra $\mathfrak{g}$, then its multiplication table is represented by a third-order structure tensor $\gamma^i_{j,k}$, whose indices i, j, k range from 1 to ℓ:

$$[e^j, e^k] = \sum_{i=1}^{\ell} \gamma^i_{j,k} e^i; \tag{B.1}$$

$$\gamma^i_{j,k} = -\gamma^i_{k,j} \text{ and } \sum_{k=1}^{\ell} \left(\gamma^h_{i,k}\gamma^k_{j,l} + \gamma^h_{j,k}\gamma^k_{l,i} + \gamma^h_{l,k}\gamma^k_{i,j}\right) = 0 \tag{B.2}$$

from the identities (Jac.1–Jac.3). It is invariant under Lie algebra isomorphisms. One way to specify a Lie algebra is to give a linear space L with a basis $\{e^1, \ldots, e^\ell\}$ and a structure tensor $\gamma^i_{j,k}$ that satisfies (B.2); then equipped with the Lie bracket defined by (B.1), L is a Lie algebra.

B.1.8 Adjoint Operator

For any Lie algebra $\mathfrak{g}$, to each element A corresponds a linear operator on $\mathfrak{g}$ called the adjoint operator and denoted by ad_A. It and its powers are defined for all $X \in \mathfrak{g}$ by

$$\mathrm{ad}_A(X) := [A, X]; \ \mathrm{ad}^0_A(X) := X, \ \mathrm{ad}^k_A(X) := [A, \mathrm{ad}^{k-1}_A(X)].$$

Every Lie algebra $\mathfrak{g}$ has an adjoint representation $\mathfrak{g} \to \mathrm{ad}_{\mathfrak{g}}$, a Lie algebra homomorphism defined for all A in $\mathfrak{g}$ by the mapping $A \mapsto \mathrm{ad}_A$. Treating $\mathfrak{g}$ as a linear space $\mathbb{R}^\ell$ allows the explicit representation of ad_A as an $\ell \times \ell$ matrix and of $\mathrm{ad}_{\mathfrak{g}}$ as a matrix Lie algebra.

B.2 Structure of Lie Algebras

The fragments of this extensive theory given here were chosen because they have immediate applications, but in a deeper exploration of bilinear systems much more Lie algebra theory is useful.

B.2.1 Nilpotent and Solvable Lie Algebras

If for every $X \in \mathfrak{g}$ the operator ad_X is nilpotent on $\mathfrak{g}$ (i.e., for some integer d we have $\mathrm{ad}_X^d(\mathfrak{g}) = 0$) then $\mathfrak{g}$ is called a nilpotent Lie algebra. An example is the Lie algebra $\mathfrak{n}(n, \mathbb{F})$, the strictly upper triangular $n \times n$ matrices with $[A, B] := AB - BA$.

The nil radical $\mathfrak{nr}_\mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ is the unique maximal nilpotent ideal of $\mathfrak{g}$. See Varadarajan [282, p. 237] for this version of Ado's Theorem and its proof.

Theorem B.2 (Ado). *A real or complex finite dimensional Lie algebra $\mathfrak{g}$ always possesses at least one faithful finite-dimensional representation ρ such that $\rho(X)$ is nilpotent for all $X \in \mathfrak{nr}_\mathfrak{g}$).*

The set

$$\mathfrak{g}' = \{[X, Y] \mid X, Y \in \mathfrak{g}\}$$

is called the derived subalgebra of $\mathfrak{g}$. A Lie algebra $\mathfrak{g}$ with the property that the nested sequence of derived subalgebras (called the derived series)

$$\mathfrak{g},\ \mathfrak{g}' = [\mathfrak{g},\ \mathfrak{g}],\ \mathfrak{g}^{(2)} = [\mathfrak{g}', \mathfrak{g}'], \dots$$

ends with $\{0\}$ is called solvable. The standard example of a solvable Lie algebra is $\mathfrak{t}(n, \mathbb{F})$, the upper triangular $n \times n$ matrices. A complex Lie algebra $\mathfrak{g}$ has a maximal solvable subalgebra, called a Borel subalgebra; any two Borel subalgebras of $\mathfrak{g}$ are conjugate.

The Cartan–Killing form for a given Lie algebra $\mathfrak{g}$ is the bilinear form $\chi(X, Y) := \mathrm{tr}(\mathrm{ad}_X\, \mathrm{ad}_Y)$. Use the adjoint representation on $\mathbb{F}^{\ell \times \ell}$ with respect to any basis $\mathbf{B} := \{B_1, \dots, B_\ell\}$ of $\mathfrak{g}$; given the structure tensor $\gamma^i_{j,k}$ for that basis,

$$\chi(B_i, B_j) = \sum_{p=1}^{\ell} \sum_{k=1}^{\ell} \gamma^k_{p,i} \gamma^p_{k,j}.$$

It is easy to check that for all $X, Y, Z \in \mathfrak{g}$, $\chi([X, Y], Z]) + \chi(X, [Y, Z]) = 0$. One of the many applications of the Cartan–Killing form is Theorem B.4 below.

The following theorem of Lie can be found, for instance, as Varadarajan [282, Th. 3.7.3]. The corollary is [282, Cor. 3.18.11].

Theorem B.3 (S. Lie). *If $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{C})$ is solvable then there exists a basis for $\mathbb{C}^n$ with respect to which all the matrices in $\mathfrak{g}$ are upper (lower) triangular.*

Corollary B.1. *Let $\mathfrak{g}$ be a solvable Lie algebra over $\mathbb{R}$. Then we can find subalgebras $\mathfrak{g}_1 = \mathfrak{g}, \mathfrak{g}_2, \dots, \mathfrak{g}_{k+1} = 0$ such that (i) $\mathfrak{g}_{i+1} \subset \mathfrak{g}_i$ as an ideal, $i \in 1, \dots, k$, and (ii)* $\dim(\mathfrak{g}_i/\mathfrak{g}_{i+1}) = 1$.

(For Lie algebras over $\mathbb{C}$ the $\mathfrak{g}_i$ can be chosen to be ideals in $\mathfrak{g}$ itself.)

B.2.2 Semisimple Lie Algebras

Any Lie algebra $\mathfrak{g}$ over $\mathbb{F}$ has a well-defined radical $\sqrt{\mathfrak{g}} \subset \mathfrak{g}$ (see [282, p. 204]); $\sqrt{\mathfrak{g}}$ is the unique solvable ideal in $\mathfrak{g}$ that is maximal in the sense that it contains all other solvable ideals. Two extreme possibilities are that $\sqrt{\mathfrak{g}} = \mathfrak{g}$, equivalent to $\mathfrak{g}$ being solvable; and that $\sqrt{\mathfrak{g}} = 0$, in which case $\mathfrak{g}$ is called semisimple. These two extremes are characterized by the following criteria; the proof is in most Lie algebra textbooks, such as [143, III.4] or [282, 3.9].

Theorem B.4 (Cartan's Criterion). *Let $\mathfrak{g}$ be a Lie subalgebra of $\mathfrak{gl}(n, \mathbb{F})$ with basis $\{B_1, \dots, B_\ell\}$.*
(i) $\mathfrak{g}$ is solvable $\Leftrightarrow \chi(B_i, B_j) = 0$ for all $B_i, B_j \in \mathbf{B}$.
(ii) $\mathfrak{g}$ is semisimple $\Leftrightarrow \chi(B_i, B_j)$ is nonsingular for all $B_i, B_j \in \mathbf{B}$.

Theorem B.5 (Levi[2]). *If $\mathfrak{g}$ is a finite-dimensional Lie algebra over $\mathbb{F}$ with radical $\mathfrak{r}$ then there exists a semisimple Lie subalgebra $\mathfrak{s} \subset \mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{s}$.*

Any two such semisimple subalgebras are conjugate, using transformations from $\exp(\mathfrak{r})$ [143, p. 92].

If $\dim \mathfrak{g} > 1$ and its only ideals are 0 and $\mathfrak{g}$ then $\mathfrak{g}$ is called simple; a semisimple Lie algebra with no proper ideals is simple. For example, $\mathfrak{so}(n)$ and $\mathfrak{so}(p, q)$ are simple.

Given a Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$, its centralizer in $\mathfrak{g}$ is

$$\mathfrak{z} := \{X \in \mathfrak{g} \,|\, [X, Y] = 0 \text{ for all } Y \in \mathfrak{h}\}.$$

Since it satisfies Jac.1–Jac.3, $\mathfrak{z}$ is a Lie algebra. The center of $\mathfrak{g}$ is the Lie subalgebra

$$\mathfrak{c} := \{X \in \mathfrak{g} \,|\, [X, Y] = 0 \text{ for all } Y \in \mathfrak{g}\}.$$

A Lie algebra is called reductive if its radical coincides with its center; in that case $\mathfrak{g} = \mathfrak{c} \oplus \mathfrak{g}'$. By [282, Th. 3.16.3] these three statements are equivalent: (i) $\mathfrak{g}$ is reductive, (ii) $\mathfrak{g}$ has a faithful semisimple representation, (iii) $\mathfrak{g}'$ is semisimple. Reductive Lie algebras are used in Appendix D.

B.3 Mappings and Manifolds

Let $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$, $W \subset \mathbb{R}^k$ be neighborhoods on which mappings $f : U \to V$ and $g : V \to W$ are respectively defined. A mapping $h : U \to W$ is called the composition of f and g, written $h = g \circ f$, if $h(x) = g(f(x))$ on U.

If the mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ has a continuous differential mapping, linear in y,

[2] For Levi's theorem see Jacobson [143, p. 91].

$$Df(x;y) := \lim_{h\to 0} \frac{f(x+hy)-f(x)}{h} = \frac{\partial f(x)}{\partial x} y,$$

it is said to be differentiable of class C^1; $f_*(x) = \frac{\partial f(x)}{\partial x}$ is known as the Jacobian matrix of f at x.

A mapping $f : U \to V$ is called C^r if it is r-times differentiable; f is smooth (C^∞) if $r = \infty$; it is real-analytic (C^ω) if each of the m components of f has a Taylor series at each $x \in \mathbb{R}^n$ with a positive radius of convergence.

Note: *If f is real-analytic on $\mathbb{R}^n$ and vanishes on any open set, then f vanishes identically on $\mathbb{R}^n$.*

Definition B.1. If $m = n$ and $\Phi : U \to V$ is one-to-one, onto V and both Φ and its inverse are C^r ($r \geq 1$) mappings[3] we say U and V are diffeomorphic and that Φ is a diffeomorphism, specifically a C^r, smooth (if $r = \infty$) or real-analytic (if $r = \omega$) diffeomorphism. If $r = 0$ then Φ (one-to-one, onto, bicontinuous) is called a homeomorphism. △

Theorem B.6 (Rank Theorem).[4] *Let $A_0 \subset \mathbb{R}^n, B_0 \subset \mathbb{R}^m$ be open sets, $f : A_0 \to B_0$ a C^r map, and suppose the rank of f on A_0 to be equal to k. If $a \in A_0$ and $b = f(a)$, then there exist open sets $A \subset A_0$ and $B \subset B_0$ with $a \in A$ and $b \in B$; open sets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$; and C^r diffeomorphisms $g : A \to U$, $h : B \to V$ such that $h \circ f \circ g^{-1}(U) \subset V$ and such that this map has the standard form $h \circ f \circ g^{-1}(x_1, \ldots, x_n) = (x_1, \ldots, x_k, 0, \ldots, 0)$.*

B.3.1 Manifolds

A differentiable manifold is, roughly, a topological space that everywhere looks locally like $\mathbb{R}^n$; for instance, an open connected subset of $\mathbb{R}^n$ is a manifold. A good example comes from geography, where in the absence of a world globe you can use an atlas: a book of overlapping Mercator charts that contains every place in the world.

Definition B.2. A real-analytic manifold $\mathcal{M}^n$ is an n-dimensional second countable Hausdorff space equipped with a real-analytic structure: that means it is covered by an atlas $\mathbf{U}_A$ of Euclidean coordinate charts related by real-analytic mappings; in symbols,

$$\mathbf{U}_A := \{(\mathbf{U}_\alpha, x_\alpha)\}_{\alpha\in\mathbb{N}} \text{ and } \mathcal{M}^n = \bigcup_{\alpha\in\mathbb{N}} \mathbf{U}_\alpha; \tag{B.3}$$

on the charts $\mathbf{U}_\alpha$ the local coordinate maps $\mathbf{U}_\alpha \to \mathbb{R}^n : \; p \mapsto x_\alpha(p)$ are continuous, and on each intersection $\mathbf{U}_\alpha \cap \mathbf{U}_\beta$ the map

[3] See Theorem B.8.

[4] See Boothby, [29, II,Th. 7.1] or Conlon [64, Th. 2.4.4].

$$x_{(\alpha\beta)} = x_{(\beta)}x_{(\alpha)}^{-1} : \; x_{(\alpha)}(\mathbf{U}_\alpha) \mapsto x_{(\beta)}(\mathbf{U}_\beta)$$

is real-analytic.[5] △

A function $f : \mathcal{M}^n \to \mathbb{R}$ is called C^ω on the real-analytic manifold $\mathcal{M}^n$ if, on every $\mathbf{U}_\alpha$ in $\mathbf{U}_A$, it is real-analytic in the coordinates $x_{(\alpha)}$. A mapping between manifolds $\Phi : \mathcal{M}^m \to N^n$ will be called C^ω if for each function $\phi \in C^\omega(N^n)$ the composition mapping $p \mapsto \Phi(p) \mapsto \phi(\Phi(p))$ is a C^ω function on $\mathcal{M}^m$. In the local coordinates $x(p)$, the rank of Φ at p is defined as the rank of the matrix

$$\Phi_*(p) := \frac{\partial\phi(x)}{\partial x}|_{x(p)},$$

and is independent of the coordinate system. If rank $\Phi_*(p) = m$ we say Φ has full rank at p; see Theorem B.6.

Definition B.1 can now be generalized. If $\Phi : \mathcal{M}^n \to N^n$ is one-to-one, onto N^n and both Φ and its inverse are C^ω we say $\mathcal{M}^n$ and N^n are diffeomorphic as manifolds and that Φ is a C^ω diffeomorphism.

Definition B.3. A differentiable map $\phi : \mathcal{M}^m \to \mathcal{N}^n$ between manifolds, $m \le n$, is called an immersion if its rank is m everywhere on $\mathcal{M}^m$. Such a map is one-to-one locally, but not necessarily globally. An immersion is called an embedding if it is a homeomorphism onto $\phi(\mathcal{N}^n)$ in the relative topology.[6]If $F : \mathcal{M} \to N$ is a one-to-one immersion and $\mathcal{M}$ is compact, then F is an embedding and $F(\mathcal{M})$ is an embedded manifold (as in [29].) △

Definition B.4. A subset $\mathcal{S}^k \subset \mathcal{M}^n$ is called a submanifold if it is a manifold and its inclusion mapping $i : \mathcal{S}^k \hookrightarrow \mathcal{M}^n$ is an immersion. It is called regularly embedded if the immersion is an embedding; its topology is the subspace topology inherited from $\mathcal{M}^n$. △

If the submanifold $\mathcal{S}^k$ is regularly embedded then at each $p \in \mathcal{S}^k$ the intersection of each of its sufficiently small neighborhoods in $\mathcal{M}^n$ with $\mathcal{S}^k$ is a single neighborhood in $\mathcal{S}^k$. An open neighborhood in $\mathcal{M}^n$ is a regularly embedded submanifold ($k = n$) of $\mathcal{M}^n$. A dense orbit on a torus (see Proposition 2.5) is a submanifold of the torus ($k = 1$) but is not regularly embedded.

Theorem B.7 (Whitney[7]). *Given a polynomial map $f : \mathbb{R}^n \to \mathbb{R}^m$, where $m \le n$, let $V = \{x \in \mathbb{R}^n | \, f(x) = 0\}$; if the maximal rank of f on V is m, the set $V_0 \subset V$ on which the rank equals m is a closed regularly embedded C^ω submanifold of $\mathbb{R}^n$ of dimension $n - m$.*

---

[5] There were many ways of defining coordinates, so $\mathcal{M}^n$ as defined abstractly in Definition B.2 is unique up to a diffeomorphism of manifolds.

[6] Sternberg [252, II.2] and Boothby [31, Ch. III] have good discussions of differentiable maps and manifolds.

[7] For Whitney's Theorem see Varadarajan [282, Th. 1.2.1]; it is used in the proof of Theorem B.7 (Malcev).

The following theorem is a corollary of the Rank Theorem B.6.

Theorem B.8 (Inverse Function Theorem). *If for $1 < r \leq \omega$ the mapping $\phi : \mathcal{N}^n \to \mathcal{M}^n$ is C^r and* $\operatorname{rank}\phi = n$ *at p, then there exist neighborhoods U of p and V of $\phi(p)$ such that $V = \phi(U)$ and the restriction of ϕ to the set U, $\phi|_U$, is a C^r diffeomorphism of U onto V; also* $\operatorname{rank}\phi|_U = n = \operatorname{rank}\phi^{-1}|_V$.

Remark B.1. For $1 \leq r \leq \infty$, Whitney's Embedding Theorem [252, Th. 4.4] states that a compact C^r manifold has a C^r embedding in $\mathbb{R}^{2n+1}$; for $r = \omega$ see Morrey's Theorem [214]: a compact real-analytic manifold has a C^ω embedding in $\mathbb{R}^{2n+1}$. A familiar example is the one-sided Klein bottle, which has dimension two, can be immersed in $\mathbb{R}^3$ but can be embedded in $\mathbb{R}^k$ only if $k \geq 5$. △

Definition B.5. An arc in a manifold $\mathcal{M}$ is for some finite T the image of a continuous mapping $\psi : [0, T] \mapsto \mathcal{M}$. We say $\mathcal{M}$ is arc-wise connected if for any two points $p, q \in \mathcal{M}$ there exists an arc such that $\psi(0) = p$ and $\psi(T) = q$. △

Example B.1. The map $x \mapsto x^3$ of $\mathbb{R}$ onto itself is a real-analytic homeomorphism but not a diffeomorphism. Such examples show that there are many real-analytic structures on $\mathbb{R}$. △

B.3.1.1 Vector Fields

The set of C^ω functions on $\mathcal{M}^n$, equipped with the usual point-by-point addition and multiplication of functions, constitutes a ring, denoted by $C^\omega(\mathcal{M})$. A derivation on this ring is a linear operator

$$\mathbf{f} : C^\omega(\mathcal{M}) \to C^\omega(\mathcal{M}) \text{ such that } \mathbf{f}(\phi\psi) = \psi\mathbf{f}(\phi) + \phi\mathbf{f}(\psi)$$

for all ϕ, ψ in the ring, and is called a C^ω vector field. The set of all such real-analytic vector fields will be denoted $\mathcal{V}(\mathcal{M}^n)$; it is an n-dimensional module over the ring of C^ω functions; the basis is $\{\partial/\partial x_i |\, i, j \in 1 \,..\, n\}$ on each coordinate chart.

For differentiable functions ψ on $\mathcal{M}$, including the coordinate functions $x(\cdot)$,

$$\mathbf{f}\psi(x) = \sum_1^n f_i(x)\frac{\partial\psi(x)}{\partial x}. \tag{B.4}$$

The tangent vector $\mathbf{f}_p$ is assigned at $p \in \mathcal{M}$ in a C^ω way by

$$\mathbf{f}_p := \mathbf{f}x(p) = \operatorname{col}(f_1(x(p)), \ldots, f_n(x(p))) \in T_p(\mathcal{M}), \tag{B.5}$$
$$\text{where } T_p(\mathcal{M}) := \{\mathbf{f}_p |\, \mathbf{f} \in \mathcal{V}(\mathcal{M})\},$$

called the tangent space at p, is isomorphic to $\mathbb{R}^n$. In each coordinate chart $\mathbf{U}_\alpha \subset \mathcal{M}$, (B.5) becomes the differential equation $\dot{x} = f(x)$.

B.3.1.2 Tangent Bundles

By the tangent bundle[8] of a real-analytic manifold $\mathcal{M}^n$, I mean the triple

$$\{T(\mathcal{M}), \pi, \{(\mathbf{U}_\alpha, x_\alpha)\}\} \text{ where } T := \bigsqcup_{p \in \mathcal{M}} T_p(\mathcal{M}^n)$$

(the disjoint union of the tangent spaces); π is the projection onto the base space $\pi(T_p(\mathcal{M}^n)) = p$; and $T(\mathcal{M}^n)$, as a $2n$-dimensional manifold, is given a topology and C^ω atlas $\{(\mathbf{U}_\alpha, x_\alpha)\}$ such that for each continuous vector field $\mathbf{f}$, if $p \in \mathbf{U}_\alpha$ then $\pi(\mathbf{f}(x_\alpha(p)) = p$. The tangent bundle is usually denoted by $\pi : T(\mathcal{M}^n) \to \mathcal{M}^n$, or just by $T(\mathcal{M}^n)$. Its restriction to $\mathbf{U}_A$ is $T(U)$. Over sufficiently small neighborhoods $U \subset \mathcal{M}^n$, the subbundle $T(U)$ factors, $T(U) = U \times \mathbb{R}^n$, and there exist nonzero vector fields on U which are constant in the local coordinates $x : U \to \mathbb{R}^n$. To say that a vector field $\mathbf{f}$ is a section of the tangent bundle is a way of saying that $\mathbf{f}$ is a way of defining the differential equation $\dot{x} = f(x)$ on $\mathcal{M}^n$.

B.3.1.3 Examples

For tori T^k the tangent bundle is the generalized cylinder $T^k \times \mathbb{R}^k$; thus a vector field $\mathbf{f}$ on T^1 assigns to each point θ a velocity $f(\theta)$ whose graph lies in the cylinder $T^1 \times \mathbb{R}$. The sphere S^2 has a nontrivial tangent bundle TS^2; it is a consequence that for every vector field $\mathbf{f}$ on S^2 there exists $p \in S^2$ such that $f(p) = 0$.

B.3.2 *Trajectories, Completeness*

Definition B.6. A flow on a manifold $\mathcal{M}$ is a continuous function $\phi : \mathbb{R} \times \mathcal{M} \to \mathcal{M}$ that for all $p \in \mathcal{M}$ and $s, t \in \mathbb{R}$ satisfies $\phi_t(\phi_s(p)) = \phi_{t+s}(p)$ and $\phi_0(p) = p$. If we replace $\mathbb{R}$ with $\mathbb{R}_+$ in this definition then ϕ is called a semi-flow. △

We will take for granted a standard[9] existence theorem: let $\mathbf{f}$ be a C^ω vector field on a C^ω manifold $\mathcal{M}^n$. For each $\xi \in \mathcal{M}^n$ there exists a $\delta > 0$ such that a unique trajectory $\{x(t) = f_t(\xi) | \, t \in (-\delta, \delta)\}$ satisfies the initial value problem

[8] Full treatments of tangent bundles can be found in [29, 64].

[9] See Sontag [249, C.3.12] for one version of the existence theorem.

$\dot{x} = f(x),\ x(0) = \xi$.[10] On $(-\delta, \delta)$ the mapping f is C^ω in t and ξ. It can be extended by composition (concatenation of arcs) $f_{t+s}(\xi) = f_t \circ f_s$ to some maximal time interval $(-T, T)$, perhaps infinite. If $T = \infty$ for all $\xi \in \mathcal{M}^n$ the vector field is said to be complete and the flow $f : \mathbb{R} \times \mathcal{M}^n \to \mathcal{M}^n$ is well defined. If $\mathcal{M}^n$ is compact, all its vector fields are complete. If $T < \infty$ the vector field is said to have finite escape time. The trajectory through ξ of the vector field $f(x) := x^2$ on $\mathcal{M}^n = \mathbb{R}$ is $\xi/(t\xi - 1)$, which leaves $\mathbb{R}$ at $t = 1/\xi$. In such a case the flow mapping is only partially defined. As the example suggests, for completeness of f it suffices that there exist positive constants c, k such that

$$\|f(x) - f(\xi)\| < c + k\|x - \xi\| \text{ for all } x, \xi.$$

However, such a growth (Lipschitz) condition is not necessary; the vector field

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

is complete; its trajectories are periodic. The composition of C^ω mappings is C^ω, so the endpoint map

$$f_t(\cdot) : \mathcal{M}^n \to \mathcal{M}^n,\ \xi \mapsto f_{t+s}(\xi)$$

is real-analytic wherever it is defined. The orbit[11] through ξ of vector field $\mathbf{f}$ is the set $\{f_t(\xi) | t \in \mathbb{R}\} \subset \mathcal{M}^n$. The set $\{f_t(\xi) | t \in \mathbb{R}_+\}$ is called a semiorbit of $\mathbf{f}$.

B.3.3 Dynamical Systems

A dynamical system on a manifold here is a triple $(\mathcal{M}, \mathcal{T}, \Theta)$, whose objects are respectively a manifold $\mathcal{M}$, a linearly ordered time-set $\mathcal{T}$, which for us will be either $\mathbb{R}_+$ or $\mathbb{Z}_+$ (in dynamical system theory the time-set is usually a group, but for our work a semigroup is preferable) and the $\mathcal{T}$-indexed transition mapping $\Theta_t : \mathcal{M} \to \mathcal{M},$ which if $\mathcal{T} = \mathbb{R}_+$ is assumed to be a semi-flow. In that case usually $\Theta_t(p) := f_t(p)$, an endpoint map for a vector field $\mathbf{f}$. Typical choices for $\mathcal{M}$ are C^ω manifolds such as $\mathbb{R}^n_*$, hypersurfaces, spheres S^n, or tori T^n.

---

[10] The Rectification Theorem (page 188) shows that if $f(\xi) = b \neq 0$, there exist coordinates in which for small $|t|$ the trajectory is $f_t(0) = tb$.

[11] What is called an orbit here and in such works as Haynes and Hermes [119] is called a *path* by those who work on path planning.

B.3.4 Control Systems as Polysystems

At this point we can define what is meant by a dynamical polysystem on a manifold. Loosely, in the definition of a dynamical system we redefine $\mathcal{F}$ as a family of vector fields, provided with a method of choosing one of them at a time in a piecewise-constant way to generate transition mappings. This concept was introduced in Lobry's thesis; see his paper [190].

In detail: given a differentiable manifold $\mathcal{M}^n$, let $\mathrm{F} \subset T(\mathcal{M}^n)$ be a family of differentiable vector fields on $\mathcal{M}^n$, for instance $\mathrm{F} := \{\mathbf{f}_\alpha\}$, each of which is supposed to be complete (see Section B.3.2), with the prescription that at any time t one may choose an α and a time interval $[t, \tau)$ on which $\dot{x} = f_\alpha(x)$. To obtain trajectories $\boldsymbol{\pi}$ for the polysystem, trajectory arcs of the vector fields $\mathbf{f}_\alpha$ are concatenated. For example, let $\mathrm{F} = \{\mathbf{f}, \mathbf{g}\}$ and $x_0 = \xi$. Choose $\mathbf{f}$ to generate the first arc $\boldsymbol{f}_t$ of the trajectory, $0 \le t < \tau$ and $\mathbf{g}$ for the second arc $\boldsymbol{g}_t$, $\tau \le t < T$. The resulting trajectory of the polysystem is

$$\boldsymbol{\pi}_t(\xi) = \begin{cases} \boldsymbol{f}_t(\xi), & t \in [0, \tau), \\ \boldsymbol{g}_{t-\tau} \circ \boldsymbol{f}_\tau(\xi), & t \in [\tau, T]. \end{cases}$$

Jurdjevic [147] has an introduction to dynamical polysystems; for the earliest work see Hermann [122], Lobry [189], and Sussmann [258–260, 262, 269].

B.3.5 Vector Field Lie Brackets

A Lie bracket is defined for vector fields (as in [147, Ch.2]) by

$$[\mathbf{f}, \mathbf{g}]\psi := \mathbf{g}(\mathbf{f}\,\psi) - \mathbf{f}(\mathbf{g}\,\psi) \ \text{ for all } \psi \in C^\omega(\mathcal{M}^n);$$

$[\mathbf{f}, \mathbf{g}]$ is again a vector field because the second-order partial derivatives of ψ cancel out.[12] Using the relation (B.5) between a vector field $\mathbf{f}$ and the corresponding $f(x) \in \mathbb{R}^n$, we can define $[f, g] := g_*(x)f(x) - f_*(x)g(x)$. The ith component of $[f, g](x)$ is

$$[f, g]_i(x) = \sum_{j=1}^{n} \left(\frac{\partial f_i(x)}{\partial x_j} g_j(x) - \frac{\partial g_i(x)}{\partial x_j} f_j(x) \right), \ i = 1, \dots, n. \tag{B.6}$$

Equipped with this Lie bracket the linear space $\mathcal{V}(\mathcal{M}^n)$ of all C^ω vector fields on $\mathcal{M}^n$ is a Lie algebra over $\mathbb{R}$. Theproperties Jac.1–Jac.3 of Section B.1 are

[12] If the vector fields are only C^1 (once differentiable) the approximation $[\mathbf{f}, \mathbf{g}] \approx \boldsymbol{f}_t\,\mathbf{g}\,\boldsymbol{f}_{-t}$ can be used, analogous to (2.3); see Chow [59] and Sussmann [258].

easily verified. $\mathcal{V}(\mathcal{M}^n)$ is infinite-dimensional; for instance, for $n = 1$ let $\mathbf{f} := x^2\frac{\partial}{\partial x}$, $\mathbf{g} := x^3\frac{\partial}{\partial x}$; then $[\mathbf{f},\mathbf{g}] = x^4\frac{\partial}{\partial x}$ and so forth.

For a linear dynamical system $\dot{x} = Ax$ on $\mathbb{R}^n$, the corresponding vector field and Lie bracket (defined as in (B.6)) are

$$\mathbf{a} := \sum_{i=1,j=1}^{n} A_{i,j}x_j\frac{\partial}{\partial x_i} = x^{\top}A^{\top}\frac{\partial}{\partial x} \text{ and } [\mathbf{a},\mathbf{b}]\psi = x^{\top}(AB - BA)^{\top}\frac{\partial\psi}{\partial x}.$$

With that in mind, $\dot{x} = [\mathbf{a},\mathbf{b}]x$ corresponds to $\dot{x} = (AB - BA)x$. This correspondence is an isomorphism of the matrix Lie algebra $\{A,B\}_{\mathcal{L}}$ with the Lie algebra of vector fields $\{\mathbf{a},\mathbf{b}\}_{\mathcal{L}}$:

$$\mathfrak{g} := \{x^{\top}A^{\top}\frac{\partial}{\partial x} \,|\, A \in \mathfrak{g}\}.$$

Affine vector fields: the Lie bracket of the affine vector fields

$$f(x) := Ax + a, \quad g(x) := Bx + b, \text{ is } [f,g](x) = [A,B]x + Ab - Ba$$

which is again affine (not linear).

Definition B.7. Given a C^{ω} map $\phi : M^m \to N^n$, let $\mathbf{f}$ and $\mathbf{f}'$ be vector fields on M^m and N^n, respectively; $\mathbf{f}$ and $\mathbf{f}'$ are called ϕ-related if they satisfy the intertwining equation

$$(\mathbf{f}'\psi) \circ \phi = \mathbf{f}(\psi \circ \phi) \tag{B.7}$$

for all $\psi \in C^{\omega}(N^n)$. △

That is, the time derivatives of ψ on $\mathcal{N}^n$ and its pullback $\psi \circ \phi$ on M^m are the same along the trajectories of the corresponding vector fields $\mathbf{f}$ and $\mathbf{f}'$. Again, at each $p \in \mathcal{M}^m$, $\mathbf{f}'_{\phi(p)} = D\phi_p \cdot \mathbf{f}_p$. This relation is[13] inherited by Lie brackets: $[\mathbf{f}',\mathbf{g}'] = D\phi \cdot [\mathbf{f},\mathbf{g}]$. In this way, $D\phi$ induces a homomorphism of Lie algebras.

B.4 Groups

A group is a set $\mathbf{G}$ equipped with:

(i) an associative composition, $\mathbf{G} \times \mathbf{G} \mapsto \mathbf{G} : (\alpha,\beta) \mapsto \alpha\beta$;
(ii) a unique identity element $\iota : \iota\alpha = \alpha = \alpha\iota$;
(iii) and for each $\alpha \in \mathbf{G}$, a unique inverse α^{-1} such that $\alpha^{-1}\alpha = \iota$ and $\alpha\alpha^{-1} = \iota$.

A subset $\mathbf{H} \subset \mathbf{G}$ is called a subgroup of $\mathbf{G}$ if $\mathbf{H}$ is closed under the group operation and contains ι. If $\alpha\mathbf{H} = \mathbf{H}\alpha$ for all $\alpha \in \mathbf{G}$, then $\mathbf{H}$ is called a normal

[13] See Conlon [64, Prop. 4.2.10].

subgroup of $\mathbf{G}$ and $\mathbf{G}/\mathbf{H}$ is a subgroup of $\mathbf{G}$. If $\mathbf{G}$ has no normal subgroups other than $\{\iota\}$ and $\mathbf{G}$ itself, it is called simple.

Definition B.8. A group homomorphism is a mapping $\pi : \mathbf{G}_1 \to \mathbf{G}_2$ such that $\iota_1 \mapsto \iota_2$ and the group operations are preserved. If one-to-one, π is a group isomorphism. △

B.4.1 Topological Groups

A topological group $\mathbf{G}$ is a topological space which is also a group for which composition ($\alpha\beta$) and inverse (α^{-1}) are continuous maps in the given topology. For example, the group $\mathrm{GL}(n, \mathbb{R})$ of all invertible $n \times n$ matrices inherits a topology as an open subset of the Euclidean space $\mathbb{R}^{n\times n}$; matrix multiplication and inversion are continuous, so $\mathrm{GL}(n, \mathbb{R})$ is a topological group. Following Chevalley [58, II.II], top($\mathbf{G}$) can be defined axiomatically by using a neighborhood base $\mathcal{W}$ of sets V containing ι and satisfying six conditions:

1. $V_1, V_2 \in \mathcal{W} \Rightarrow V_1 \cap V_2 \in \mathcal{W}$;
2. $\bigcap\{V \in \mathcal{W}\} = \{\iota\}$;
3. if $V \subset U \subset \mathbf{G}$ and $V \in \mathcal{W}$ then $U \in \mathcal{W}$;
4. if $V \in \mathcal{W}$, then $\exists V_1 \in \mathcal{W}$ s.t. $V_1 V \in \mathcal{W}$;
5. $\{V |\, V^{-1} \in \mathcal{W}\} = \mathcal{W}$;
6. if $\sigma \in \mathcal{W}$ then $\{\sigma V \sigma^{-1} |\, V \in \mathcal{W}\} = \mathcal{W}$.

Then neighborhood bases of all other points in $\mathbf{G}$ can be obtained by right or left actions of elements of $\mathbf{G}$.

Definition B.9. If $\mathbf{G}$, $\mathbf{H}$ are groups and $\mathbf{H}$ is a subgroup of $\mathbf{G}$ then the equivalence relation $\sim$ on $\mathbf{G}$ defined by $\alpha \sim \beta$ if and only if $\alpha^{-1}\beta \in \mathbf{H}$ partitions $\mathbf{G}$ into classes (left cosets); α is a representative of the coset $\{\alpha\mathbf{H}\}$. The cosets compose by the rule $(\alpha\mathbf{H})(\beta\mathbf{H}) = \alpha\beta\mathbf{H}$, and the set of classes $\mathbf{G}/\mathbf{H} := \mathbf{G}/\sim_R$ is called a (left) coset space. △

Definition B.10. Let $\mathbf{G}$ be a group and S a set. Then $\mathbf{G}$ is said to act on S if there is a mapping $F : \mathbf{G} \times S \to S$ satisfying, for all $x \in S$,

(i) $F(\iota, x) = x$ and
(ii) if $\alpha, \beta \in \mathbf{G}$ then $F(\alpha, F(\beta, x)) = F(\alpha\beta, x)$.

Usually one writes $F(\alpha, x)$ as αx. If $\mathbf{G}$ is a topological group and S has a topology, the action is called continuous if F is continuous. One says $\mathbf{G}$ acts transitively on S if for any two elements $x, y \in S$ there exists $\gamma \in \mathbf{G}$ such that $\gamma x = y$. A group, by definition, acts transitively on itself. △

B.5 Lie Groups

Definition B.11. A real Lie group is a topological group $\mathbf{G}$ (the identity element is denoted by ι) that is also a finite-dimensional C^ω manifold;[14] the group's composition $\rho, \sigma \mapsto \rho\sigma$ and inverse $\rho \mapsto \sigma : \rho\sigma = \iota$ are both real-analytic. △

A connected Lie group is called an analytic group. Complex Lie groups are defined analogously. Like any other C^ω manifold, a Lie group $\mathbf{G}$ has a tangent space $T_p(\mathbf{G})$ at each point p.

Definition B.12 (Translations). To every element α in a Lie group $\mathbf{G}$ there correspond two mappings defined for all $\sigma \in \mathbf{G}$ — the right translation $R_\rho : \sigma \mapsto \sigma\alpha$ and the left translation $L_\alpha : \sigma \mapsto \alpha\sigma$. If $\rho \in \mathbf{G}$ then R_ρ and L_ρ are real-analytic diffeomorphisms. Furthermore, $\mathbf{G} \to R_\rho\mathbf{G}$ and $\mathbf{G} \to L_\rho\mathbf{G}$ are group isomorphisms. △

A vector field $\mathbf{f} \in \mathcal{V}(\mathbf{G})$ is called right-invariant if $DR_\rho\mathbf{f} = \mathbf{f}$ for each $\rho \in \mathbf{G}$, left-invariant if $DL_\rho\mathbf{f} = \mathbf{f}$. The bracket of two right-invariant (left-invariant) vector fields is right-invariant (left-invariant). We will use right-invariance for convenience.

Definition B.13. The (vector field) Lie algebra of a Lie group $\mathbf{G}$, denoted as $\mathcal{L}(\mathbf{G})$, is the linear space of *right-invariant* vector fields on $\mathbf{G}$ equipped with the vector field Lie bracket of Section B.3.5 (see Varadarajan [282, p. 51]). For an equivalent definition see Section B.5.2. △

B.5.1 Representations and Realizations

A representation of a Lie group $\mathbf{G}$ is an analytic (or C^ω if $\mathbb{F} = \mathbb{R}$) isomorphism $\varrho : \mathbf{G} \to \mathbf{H}$ where for some n, $\mathbf{H}$ is a subgroup of $\mathrm{GL}(n, \mathbb{F})$. A one-dimensional representation $\varrho : \mathbf{G} \to \mathbb{F}$ is called a character of the group. One representation of the translation group $\mathbb{R}^2$, for example, is

$$\varrho(\mathbb{R}^2) = \begin{bmatrix} 1 & x_1 & x_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \subset \mathrm{GL}(3, \mathbb{R}).$$

In this standard definition of representation, the action of the group is linear (matrix multiplication).

A group realization, sometimes called a representation, is an isomorphism through an action (not necessarily linear) of a group on a linear space, for

[14] In the definition of Lie group C^ω can be replaced by C^r where r can be as small as 0; if so, it is then a theorem (see Montgomery and Zippin [213]) that the topological group is a C^ω manifold.

instance a matrix $T \in \mathrm{GL}(n, \mathbb{R})$ can act on the matrices $X \in \mathbb{F}^{n \times n}$ by $(T, X) \mapsto T^{-1}XT$.

B.5.2 The Lie Algebra of a Lie Group

The tangent space at the identity $T_\iota(\mathbf{G})$ can be given the structure of a Lie algebra $\mathfrak{g}$ by associating each tangent vector Z with the unique right-invariant vector field $\mathbf{z}$ such that $\mathbf{z}(\iota) = Z$ and defining Lie brackets by $[Y, Z] = [\mathbf{y}, \mathbf{z}](\iota)$. The dimension ℓ of $\mathfrak{g}$ is that of $\mathbf{G}$ and it is, as a Lie algebra, the same as $\mathcal{L}(\mathbf{G})$, the *Lie algebra of* $\mathbf{G}$ defined in Definition B.13. The tangent bundle $T(\mathbf{G})$ (see Section B.3.1.2) assigns a copy of the space $\mathfrak{g}$ at $\iota \in \mathbf{G}$ and at each point $X \in \mathbf{G}$ assigns the space $X^{-1}\mathfrak{g}X$. Theorem B.10, below, states that the subset of $\mathbf{G}$ connected to its identity element (identity component) is uniquely determined by $\mathfrak{g}$.

If all of its elements commute, $\mathbf{G}$ is called an Abelian group. On the Abelian Lie groups $\mathbb{R}^n \oplus T^p$ the only right-invariant vector fields are the constants $\dot{x} = a$, which are also left-invariant.

If $\mathbf{G}$ is a *matrix* Lie group containing the matrices X and S then $R_S X = XS$; $DR_S = S$ acting on its left. If A is a tangent vector at I for $\mathbf{G}$ then the vector field $\mathbf{a}$ given by $\dot{X} = AX$ is right-invariant on $\mathbf{G}$: for any right translation $Y = XS$ by an element $S \in \mathbf{G}$, one has $\dot{Y} = AY$.

In the special case $\mathbf{G} = \mathrm{GL}(n, \mathbb{R})$, the differential operator corresponding to $\dot{Y} = AY$ is (using the coordinates $x_{i,j}$ of $X \in \mathrm{GL}(n, \mathbb{R})$)

$$\mathbf{a} = \sum_{i,j,k=1}^{n} a_{i,k} x_{k,j} \frac{\partial}{\partial x_{i,j}}.$$

Left translation and left-invariant vector fields are more usual in the cited textbooks[15] but translating between the two approaches is easy and right-invariance is recognized to be a convenient choice in studying dynamical systems on groups and homogeneous spaces.

B.5.3 Lie Subgroups

If the subgroup $\mathbf{H} \subset \mathbf{G}$ has a manifold structure for which the inclusion map $i : \mathbf{H} \hookrightarrow \mathbf{G}$ is both a one-to-one immersion and a group homomorphism, then we will call $\mathbf{H}$ a Lie subgroup of $\mathbf{G}$. The relative topology of $\mathbf{H}$ in $\mathbf{G}$ may not coincide with its topology as a Lie group; the dense line on the torus T^2 (Example 2.5) is homeomorphic and isomorphic, as a group, to $\mathbb{R}$ but any

[15] Cf. Jurdjevic [147].

neighborhood of it on the torus has an infinity of preimages, so i is not a regular embedding.

Definition B.14. A matrix Lie group is a Lie subgroup of $GL^+(n, \mathbb{R})$. △

Example B.2. The right- and left-invariant vector fields on matrix Lie groups[16] are related by the differentiable map $X \mapsto X^{-1}$: if $\dot{X} = AX$ let $Z = X^{-1}$, then $\dot{Z} = -ZA$. △

Lemma B.1. *If* $\mathbf{G}$ *is a Lie group and* $\mathbf{H} \subseteq \mathbf{G}$ *is a* **closed** *(in the topology of* $\mathbf{G}$*) subset that is also an abstract subgroup, then* $\mathbf{H}$ *is a properly embedded Lie subgroup [64, Th. 5.3.2].*

Theorem B.9. *If* $\mathbf{H} \subset \mathbf{G}$ *is a normal subgroup of* $\mathbf{G}$ *then the cosets of* $\mathbf{H}$ *constitute a Lie group* $K := \mathbf{G}/\mathbf{H}$.

An element of K (a coset of $\mathbf{H}$) is an equivalence class $\{\mathbf{H}\sigma | \sigma \in \mathbf{G}\}$. The quotient map takes $\mathbf{H}$ to $\iota \in K$; by normality it takes a product $\sigma_1\sigma_2$ to $\mathbf{H}(\sigma_1\sigma_2)$, so $G \to K$ is a group homomorphism and K can be identified with a closed subgroup of $\mathbf{G}$ which by the preceding lemma is a Lie subgroup.

Note that a Lie group is simple (page 239) if and only if its Lie algebra is simple (page 231).

B.5.4 Real Algebraic Groups

By a real algebraic group $\mathbf{G}$ is meant a subgroup of $GL(n, \mathbb{R})$ that is the zero-set of a finite set of real polynomial functions. If $\mathbf{G}$ is a real algebraic group then it is a closed Lie subgroup of $GL(n, \mathbb{R})$ [282, Th. 2.1.2]. (The rank of the map is preserved under right translations by elements of the subgroup, so it is constant on $\mathbf{G}$. Use the Whitney Theorem B.7.)

B.5.5 Lie Group Actions and Orbits

Corresponding to the general notion of action (Definition B.10), a Lie group $\mathbf{G}$ may have an action $F : \mathbf{G} \times \mathcal{M}^n \to \mathcal{M}^n$ on a real-analytic manifold $\mathcal{M}^n$, and we usually are interested in real-analytic actions. In that case $F(\alpha, \cdot)$ is a C^ω diffeomorphism on $\mathcal{M}^n$; the usual notation is $(\alpha, x) \mapsto \alpha x$.

Given $x \in \mathcal{M}^n$, the set

$$\mathbf{G}x := \{y | \, y = \alpha x, \alpha \in \mathbf{G}\} \subset \mathcal{M}^n$$

[16] See Jurdjevic [147, Ch. 2].

is called the orbit of $\mathbf{G}$ through x. (This definition is consistent with that given for orbit of a vector field's trajectory in Section B.3.2.) Note that an orbit $\mathbf{G}x$ is an equivalence class in $\mathcal{M}^n$ for the relation defined by the action of $\mathbf{G}$. For a full treatment of Lie group actions and orbits see Boothby [29].

B.5.6 Products of Lie Groups

Definition B.15 (Direct product). Given two Lie groups $\mathbf{G}_1$, $\mathbf{G}_2$ their direct product $\mathbf{G}_1 \times \mathbf{G}_2$ is the set of pairs $\{(\alpha_1, \alpha_2) | \alpha_i \in \mathbf{G}_i\}$ equipped with the product topology, the group operation $(\alpha_1, \alpha_2)(Y_1, Y_2) = (\alpha_1 Y_1, \alpha_2, Y_2)$ and identity $\iota := (\iota_1, \iota_2)$. Showing that $\mathbf{G}_1 \times \mathbf{G}_2$ is again a Lie group is a simple exercise. A simple example is the torus $T_2 = T_1 \times T_2$. Given $\alpha_i \in \mathbf{G}_i$ and their matrix representations $\varrho_1(\alpha_1) = A_1$, $\varrho_2(\alpha_2) = A_2$ on respective linear spaces $\mathbb{F}^m, \mathbb{F}^n$, the product representation of (α_1, α_2) on $\mathbb{F}^{m+n}$ is $\mathrm{diag}(A_1, A_2)$.

The Lie algebra for $\mathbf{G}_1 \times \mathbf{G}_2$ is $\mathfrak{g}_1 \oplus \mathfrak{g}_2 := \{(v_1, v_2) | v_i \in \mathfrak{g}_i$ with the Lie bracket $[(v_1, v_2), (w_1, w_2)] := ([v_1, w_1], [v_2, w_2])$. △

Definition B.16 (Semidirect product). Given a Lie group $\mathbf{G}$ with Lie algebra $\mathfrak{g}$ and a linear space V on which $\mathbf{G}$ acts, the semidirect product $V \rtimes \mathbf{G}$ is the set of pairs $\{(v, \alpha) | v \in V, \alpha \in \mathbf{G}\}$ with group operation $(v, \alpha) \cdot (w, \beta) := (v + \alpha w, \alpha\beta)$ and identity $(0, \iota)$. The Lie algebra of $V \rtimes \mathbf{G}$ is the semidirect sum $V \dotplus \mathfrak{g}$ defined as the set of pairs $\{(v, X)\}$ with the Lie bracket $[(v, X), (w, Y)] := (Xw - Yv, [X, Y])$. For the matrix representation of the Lie algebra $V \dotplus \mathfrak{g}$ see Section 3.8. △

B.5.7 Coset Spaces

Suppose $\mathbf{G}$, $\mathbf{H}$ are Lie groups such that $\mathbf{H}$ is a **closed** subgroup of $\mathbf{G}$; let $\mathfrak{g}$ and $\mathfrak{h}$ be the corresponding Lie algebras. Then (see Helgason [120, Ch. II] or Boothby [31, Ch. IV]) on $\mathbf{G}$ the relation $\sim$ defined by

$$\beta \sim \alpha : \{\exists \gamma \in \mathbf{H} | \beta = \alpha\gamma\}$$

is an equivalence relation and partitions $\mathbf{G}$; the components of the partition are $\mathbf{H}$ orbits on $\mathbf{G}$. The quotient space $\mathcal{M}^n = \mathbf{G}/\mathbf{H}$ (with the natural topology in which the projection $\mathbf{G} \to \mathbf{G}/\mathbf{H}$ is continuous and open) is a Hausdorff submanifold of $\mathbf{G}$. It is sometimes called a coset space or orbit space; like Helgason [120], Varadarajan [282] and Jurdjevic [147], I call $\mathcal{M}^n$ a homogeneous space. $\mathbf{G}$ acts transitively on $\mathcal{M}^n$: given a pair of points p, q in $\mathcal{M}^n$, there exists $\alpha \in \mathbf{G}$ such that $\alpha p = q$.

Another viewpoint of the same facts: given a locally compact Hausdorff space $\mathcal{M}^n$ on which $\mathbf{G}$ acts transitively, and any point $p \in \mathcal{M}^n$, then the subgroup $\mathbf{H}_p \subset \mathbf{G}$ such that $\alpha p = p$ for all $\alpha \in \mathbf{H}_p$ is (obviously) closed; it is called the isotropy subgroup[17] at p. The mapping $\alpha\mathbf{H} \to Xp$ is a diffeomorphism of $\mathbf{G}/\mathbf{H}$ to $\mathcal{M}^n$. The Lie algebra corresponding to $\mathbf{H}_p$ is the isotropy subalgebra

$$\mathfrak{h}_p = \{X \in \mathfrak{g} \mid Xp = 0\}.$$

B.5.8 Exponential Maps

In the theory of differentiable manifolds the notion of "exponential map" has several definitions. When the manifold is a Lie group $\mathbf{G}$ and a dynamical system is induced on $\mathbf{G}$ by a right-invariant vector field $\dot{p} = \mathbf{f}p$, $p(0) = \iota$, then the exponential map Exp is defined by the endpoint of the trajectory from ι at time t, that is, $\mathrm{Exp}(t\mathbf{f}) := p(t)$. For a matrix Lie group, Exp coincides with the matrix exponential of Section 1.1.2; if $F \in \mathfrak{gl}(n, \mathbb{R})$ and $\dot{X} = FX$, $X(0) = I$, then $\mathrm{Exp}(tF) = X(t) = e^{tF}$. For our exemplary $a \in \mathbb{R}^2$, $\dot{x} = a$, $x(0) = 0$, and $\mathrm{Exp}(at) = ta$. For an exercise, use the the matrix representation of $\mathbb{R}^2$ in $\mathrm{GL}(3, \mathbb{R})$ above to see the relation of exp to Exp.

Theorem B.10 (Correspondence). *Let $\mathbf{G}$ be a Lie group, $\mathfrak{g}$ its Lie algebra, and $\mathfrak{h}$ a subalgebra of $\mathfrak{g}$. Then there is a connected Lie subgroup $\mathbf{H}$ of $\mathbf{G}$ whose Lie algebra is $\mathfrak{h}$. The correspondence is established near the identity by $\mathrm{Exp} : \mathfrak{h} \to \mathbf{H}$, and elsewhere by the right action of $\mathbf{G}$ on itself.*

Discussion. See Boothby [29, Ch. IV, Th. 8.7]; a similar theorem [282, Th. 2.5.2] in Varadarajan's book explains, in terms of vector field distributions and maximal integral manifolds, the bijective correspondence between Lie subalgebras and connected subgroups (called analytic subgroups in [282]), that are not necessarily closed. $\triangle$

B.5.9 Yamabe's Theorem

Lemma B.2 (Lie Product Formula). *For any $A, B \in \mathbb{F}^{n\times n}$ and uniformly for $t \in [0, T]$*

$$e^{t(A+B)} = \lim_{n\to\infty} \left(\exp(\frac{tA}{n}) \exp(\frac{tB}{n}) \right)^n . \tag{B.8}$$

Discussion. For fixed t this is a consequence of (1.6) and the power series in Definition 1.3; see Varadarajan [282, Cor. 2.12.5]) for the details, which show the claimed uniformity. In the context of linear operators on Banach spaces, it is known as the Trotter formula and can be interpreted as a method for solving $\dot{X} = (A + B)X$. $\triangle$

---

[17] An isotropy subgroup $\mathbf{H}_p$ is also called a stabilizer subgroup.

Theorem B.11 (Yamabe). *Let* **G** *be a Lie group, and let* **H** *be an arc-wise connected subgroup of* **G**. *Then* **H** *is a Lie subgroup of* **G**.

Discussion. For arc-wise connectedness see Definition B.5. Part of the proof of Theorem B.11 given by M. Goto [107] can be sketched here for the case $\mathbf{G} = \mathrm{GL}^+(n, \mathbb{R})$ (matrix Lie groups).

Let **H** be an arc-wise connected subgroup of $\mathrm{GL}^+(n, \mathbb{R})$. Let $\mathfrak{h}$ be the set of $n \times n$ matrices L with the property that for any neighborhood U of the identity matrix in $\mathrm{GL}^+(n, \mathbb{R})$, there exists an arc

$$\alpha : [0, 1] \to \mathbf{H} \text{ such that } \alpha(0) = I, \ \alpha(t) \in \exp(tL)U, 0 \leq t \leq 1.$$

First show that if $X \in \mathfrak{h}$ then so does λX, $\lambda \in \mathbb{R}$. Use Lemma B.2 to show that for sufficiently small X and Y, $X + Y \in \mathfrak{h}$. The limit (2.3) shows that $X, Y \in \mathfrak{h} \Rightarrow [X, Y] \in \mathfrak{h}$. Therefore $\mathfrak{h}$ is a matrix Lie algebra. Use the Brouwer fixed point theorem to show that **H** is the connected matrix Lie subgroup of $\mathrm{GL}^+(n, \mathbb{R})$ that corresponds to $\mathfrak{h}$. △

This page intentionally left blank

Appendix C
Algebraic Geometry

The brief excursion here into algebraic geometry on affine spaces $\mathbb{F}^n$, where $\mathbb{F}$ is $\mathbb{C}$ or $\mathbb{R}$, is guided by the book of Cox et al. [67], which is a good source for basic facts about polynomial ideals, affine varieties, and symbolic algebraic computation.

C.1 Polynomials

Choose a ring of polynomials in n indeterminates $\mathbf{R} := \mathbb{F}[x_1, \dots, x_n]$ over a field $\mathbb{F}$ that contains $\mathbb{Q}$, such as $\mathbb{R}$ or $\mathbb{C}$. A subset $\mathcal{I} \subset \mathbf{R}$ is called an ideal if it contains 0 and is closed under multiplication by polynomials in $\mathbf{R}$.

A multi-index is a list $\alpha = (d_1, d_2, \dots, d_n)$ of nonnegative integers; its cardinality is $|\alpha| := d_1 + \cdots + d_n$. A monomial of degree d in $\mathbf{R}$, such as $x^\alpha := x_1^{d_1} \dots x_n^{d_n}$, is a product of indeterminates x_i such that $|\alpha| = d$.

In $\mathbf{R}$ a homogeneous polynomial p of degree d is a sum of monomials, each of degree d. For such symbolic computations as finding greatest common divisors or Gröbner bases, one needs a total order $\succ$ on monomials x^α; it must satisfy the product rule $x^\alpha x^\beta \succ x^\alpha$ for any two monomials x^α, x^β. Since the d_i are integers, to each monomial x^α corresponds a lattice point $\mathrm{col}(d_1, \dots, d_n) \in \mathbb{Z}_+^n$; so $\succ$ is induced by a total order on $\mathbb{Z}_+^n$. It will suffice here to use the lexicographic order, called *lex*, on $\mathbb{Z}_+^n$. Given positive integer multi-indices $\alpha := \mathrm{col}(\alpha_1, \dots, \alpha_n)$ and $\beta := \mathrm{col}(\beta_1, \dots, \beta_n)$ we say $\alpha \succ_{lex} \beta$ if, in the vector difference $\alpha - \beta \in \mathbb{Z}^n$, the leftmost nonzero entry is positive. Furthermore, we write $x^\alpha \succ_{lex} x^\beta$ if $\alpha \succ_{lex} \beta$.

Let $\{q_1, \dots, q_r\}$ be a finite set of linearly independent polynomials in $\mathbf{R}$; the ideal that this set generates in $\mathbf{R}$ is

$$\langle q_1, \dots, q_s \rangle := \left\{ \sum_{i=1}^{s} h_i(x) q_i(x) \,\middle|\, h_i \in \mathbf{R} \right\},$$

D.L. Elliott, *Bilinear Control Systems*, Applied Mathematical Sciences 169, DOI 10.1007/978-1-4020-9613-6_11, © Springer Science+Business Media B.V. 2009

where the bold angle brackets are used to avoid conflict with the inner product $\langle , \rangle$ notation. By *Hilbert's Basis Theorem* [67, Ch. 2, Th. 4], for any $\mathbb{F}$ every ideal $\mathcal{I} \subset \mathbf{R}$ has a finite generating set $\boldsymbol{\langle} p_1, \ldots, p_k \boldsymbol{\rangle}$, called a basis. (If $\mathcal{I}$ is defined by some linearly independent set of generators then that set is a basis.) Different bases for $\mathcal{I}$ are related in a way, called the method of Gröbner bases, analogous to gaussian elimination for linear equations. The bases of $\mathcal{I}$ are an equivalence class; they correspond to a single geometric object, an affine variety, to be defined in the next section.

C.2 Affine Varieties and Ideals

The affine variety V defined by a list of polynomials $\{q_1, \ldots, q_r\}$ in $\mathbf{R}$ is the set in $\mathbb{F}^n$ where they all vanish; this sets up a mapping $\mathbf{V}$ defined by

$$\mathbf{V}\boldsymbol{\langle} q_1, \ldots, q_r \boldsymbol{\rangle} := \{x \in \mathbb{F}^n \,|\, q_i(x) = 0,\ 1 \le i \le r\}\,.$$

The more polynomials in the list, the smaller the affine variety. The affine varieties we will deal with in this chapter are defined by homogeneous polynomials, so they are cones. (A real cone is a set $S \subset \mathbb{R}^n$ such that $\mathbb{R}S = S$.)

Given an affine variety $V \subset \mathbb{F}^n$ then $\mathcal{I} := \mathbf{I}(V)$ is the ideal generated by all the polynomials that vanish on V. Unless $\mathbb{F} = \mathbb{C}$ the correspondence between ideals and varieties is not one-to-one.

C.2.1 Radical Ideals

Given a single polynomial $q \in \mathbf{R}$, possibly with repeated factors, the product $\tilde{q}$ of its *prime factors* is called the reduced form or square-free form of q; clearly $\tilde{q}$ has the same zero-set as q. The algorithm (C.1) we can use to obtain $\tilde{q}$ is given in [67, Ch. 4, §2 Prop. 12]:

$$q_r = \frac{q}{\mathrm{GCD}\left(q, \dfrac{\partial q}{\partial x_1}, \ldots, \dfrac{\partial q}{\partial x_n}\right)}. \tag{C.1}$$

It is used in an algorithm in Fig. 2.4.

An ideal is called radical if for any integer $m \ge 1$

$$f^m \in \mathcal{I} \Rightarrow f \in \mathcal{I}.$$

The radical of the ideal $\mathcal{I} := \boldsymbol{\langle} q_1, \ldots, q_s \boldsymbol{\rangle}$ is defined as

$$\sqrt{\mathcal{I}} := \{f \,|\, f^m \in \mathcal{I} \text{ for some integer } m \geq 1\}.$$

and [67, p. 174] is a radical ideal. The motive for introducing radical ideals is that over any algebraically closed field $\mathbf{k}$ (such as $\mathbb{C}$) they characterize affine varieties; that is the strong version of Hilbert's Nullensatz:

$$\mathbf{I}(\mathbf{V}(\mathcal{I})) = \sqrt{\mathcal{I}}.$$

Over $\mathbf{k}$ the mappings $\mathbf{I}(\cdot)$ from affine varieties to radical ideals and $\mathbf{V}(\cdot)$ from radical ideals to varieties are inverses of each other.

C.2.2 Real Algebraic Geometry

Some authors prefer to reserve the term *radical ideal* for use with polynomial rings over an algebraically closed field, but [67] makes clear that it can be used with $\mathbb{R}[x_1, \ldots, x_n]$. Note that in this book, like [67], in symbolic computation for algebraic geometry the matrices and polynomials have coefficients in $\mathbb{Q}$ (or extensions, e.g., $\mathbb{Q}[\sqrt{-2}]$). Mostly we deal with affine varieties in $\mathbb{R}^n$. In the investigations of Lie rank in Chapter 2 we often need to construct real affine varieties from ideals; rarely, given such a variety V, we need a corresponding real radical ideal.

Definition C.1. (See [27, Ch. 4].) Given an ideal $\mathcal{I} \subset \mathbb{R}[x_1, \ldots, x_n]$, the real radical $\sqrt[R]{\mathcal{I}}$ is the set of polynomials p in $\mathbb{R}[x_1, \ldots, x_n]$ such that there exists $m \in \mathbb{Z}_+$ and polynomials

$$q_1, \ldots, q_k \in \mathbb{R}[x_1, \ldots, x_n] \text{ such that } p^{2m} + q_1^2 + \cdots + q_k^2 \in \mathcal{I}. \qquad \triangle$$

Calculating a real radical is difficult, but for ideals generated by the homogeneous polynomials of Lie rank theory many problems yield to rules such as

$$q(x_1, \ldots, x_r) \gg 0 \;\Rightarrow \mathbf{V}\langle q, p_1, \ldots, p_s\rangle = \mathbf{V}\langle x_1, \ldots, x_r, p_1, \ldots, p_s\rangle. \tag{C.2}$$

On $\mathbb{R}^2$, for example, $\mathbf{V}\langle x_1^2 - x_1x_2 + x_2^2\rangle = \{0\}$; and on $\mathbb{R}^n$ the affine variety $\mathbf{V}\langle x_2^2 + \cdots + x_n^2\rangle$ is the x_1-axis.

For $\mathbb{R}$ there is no complete correspondence between radical ideals and (affine) varieties that would be true for $\mathbb{C}$, but (see [67, Ch. 4, §2 Th. 7]) still the following relations can be shown:

if $\mathcal{I}_1 \subset \mathcal{I}_2$ are ideals then $\mathbf{V}(\mathcal{I}_1) \supset \mathbf{V}(\mathcal{I}_2)$;

if $\mathrm{V}_1 \subset \mathrm{V}_2$ are varieties then $\mathbf{I}(\mathrm{V}_1) \supset \mathbf{I}(\mathrm{V}_2)$;

and if V is any affine variety, $\mathbf{V}(\mathbf{I}(\mathrm{V})) = \mathrm{V}$, so the mapping $\mathbf{I}$ from affine varieties to ideals is one-to-one.

C.2.3 Zariski Topology

The topology used in algebraic geometry is often the Zariski topology. It is defined by its closed sets, which are the affine varieties in $\mathbb{F}^n$. The Zariski closure of a set $S \in \mathbb{F}^n$ is denoted $\bar{S}$; it is the smallest affine variety in $\mathbb{F}^n$ containing S. We write $\bar{S} := \mathbf{V}(\mathbf{I}(S))$. The Zariski topology has weaker separation properties than the usual topology of $\mathbb{F}^n$, since it does not satisfy the Hausdorff T_2 property. It is most useful over an algebraically closed field such as $\mathbb{C}$ but is still useful, with care, in real algebraic geometry.

Appendix D
Transitive Lie Algebras

D.1 Introduction

The connected Lie subgroups of $\mathrm{Gl}(n, \mathbb{R})$ that are transitive on $\mathbb{R}^n_*$ are, loosely speaking, canonical forms for the symmetric bilinear control systems of Chapter 2 and are important in Chapter 7. Most of their classification was achieved by Boothby [32] at the beginning of our long collaboration. The corresponding Lie algebras $\mathfrak{g}$ (also called transitive because $\mathfrak{g}x = \mathbb{R}^n$ for all $x \in \mathbb{R}^n_*$) are discussed and a corrected list is given in Boothby-Wilson [32]; that work presents a rational algorithm, using the theory of semisimple Lie algebras, that determines whether a generated Lie algebra $\{A, B\}_{\mathcal{L}}$ is transitive.

Independently Kramer [162], in the context of 2-transitive groups, has given a complete list of transitive groups that includes one missing from [32], $\mathbf{Spin}(9, 1)$. The list of prehomogeneous vector spaces given, again independently, by T. Kimura [159] is closely related; it contains groups with open orbits in various *complex* linear spaces and is discussed in Section 2.10.2.

*D.1.1 Notation for the Representations**

The list of transitive Lie algebras uses some matrix representations standard in the literature: the $\boldsymbol{\alpha}$ representation of complex matrices, which is induced by the identification $\mathbb{C}^n \to \mathbb{R}^{2n}$ of (A.2), and the $\boldsymbol{\beta}$ representation of matrices over the quaternion field $\mathbb{H}$. If A, B, P, Q, R, S are real square matrices

$$\boldsymbol{\alpha}(A + iB) = \begin{bmatrix} A & B \\ -B & A \end{bmatrix}; \quad \boldsymbol{\beta}(P\mathbf{1} + Q\mathbf{i} + R\mathbf{j} + S\mathbf{k}) = \begin{bmatrix} P & -Q & -R & -S \\ Q & P & -S & R \\ R & S & P & -Q \\ S & -R & Q & P \end{bmatrix}.$$

DOI 10.1007/978-1-4020-9613-6_12,

The argument of $\boldsymbol{\alpha}$ is a complex matrix; see Example 2.27. $(\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k})$ are unit quaternions; the argument of $\boldsymbol{\beta}$ is a right quaternion matrix.

Remark D.1. $\boldsymbol{\beta}(x_1\mathbf{1}+x_2\mathbf{i}+x_3\mathbf{j}+x_4\mathbf{k}), x \in \mathbb{R}^4$, is a representation of the quaternion algebra $\mathbb{H}$; Hamilton's equations $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -\mathbf{1}$ are satisfied. The pure quaternions are those with $x_1 = 0$. $\mathbb{H}$ is a four-dimensional Lie subalgebra of $\mathfrak{gl}(4, \mathbb{R})$ (with the usual bracket $[a, b] = ab - ba$). The corresponding Lie group is $\mathbb{H}_*$; see Example 2.6 and in Section D.2 see item I.2. For this Lie algebra the singular polynomial is $\mathcal{P}(x) = (x_1^2 + x_2^2 + x_3^2 + x_4^2)^2$. △

For a representation of $\mathfrak{spin}(9)$ it is convenient to use the Kronecker product $\otimes$ (see Section A.3) and the 2×2 matrices

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{D.1}$$

To handle $\mathfrak{spin}(9, 1)$ the octonions $\mathbb{O}$ are useful; quoting Bryant [45]

"$\mathbb{O}$ is the unique $\mathbb{R}$-algebra of dimension 8 with unit $\mathbf{1} \in \mathbb{O}$ endowed with a positive definite inner product $\langle\cdot\rangle$ satisfying

$$\langle \mathbf{xy}, \mathbf{xy} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle \text{ for all } \mathbf{xy} \in \mathbb{O}.$$

As usual, the norm of an element $\mathbf{x} \in \mathbb{O}$ is denoted $|\mathbf{x}|$ and defined as the square root of $\langle \mathbf{x}, \mathbf{x} \rangle$. Left and right multiplication by $\mathbf{x} \in \mathbb{O}$ are defined by maps $L_{\mathbf{x}}, R_{\mathbf{x}} : \mathbb{O} \to \mathbb{O}$ that are isometries when $|\mathbf{x}| = 1$."

The symbol † indicates the octonion conjugate: $x^\dagger := 2\langle \mathbf{x}, \mathbf{1} \rangle \mathbf{1} - \mathbf{x}$.

*D.1.2 Finding the Transitive Groups**

The algorithmic treatment in [32] is valuable in reading [29], which will be sketched here. Boothby constructed his list using two[1] basic theorems, of independent interest.

Theorem D.1. *Suppose that* $\mathbf{G}$ *is transitive on* $\mathbb{R}^n_*$ *and* $\mathbf{K}$ *is a maximal compact subgroup of* $\mathbf{G}$*, and let* $\langle x, y \rangle_{\mathbf{K}}$ *be a* $\mathbf{K}$*-invariant inner product on* $\mathbb{R}^n$*. Then* $\mathbf{K}$ *is transitive on the unit sphere* $\tilde{S}^{n-1} := \{x \in \mathbb{R}^n | \, \langle x, x \rangle_{\mathbf{K}} = 1\}$.

Proof (Sketch). From the transitive action of $X \in \mathbf{G}$ on $\mathbb{R}^n_*$ we obtain a differentiable action $*$ of X on the standard sphere S^{n-1} as follows. For $x \in S^{n-1}$ let $X{*}x := \frac{Xx}{\|Xx\|}$, a unit vector. The existence of $*$ implies that $\mathbf{G}$ has a maximal compact subgroup $\mathbf{K}$ whose action by $*$ is transitive on S^{n-1}; and that means $\mathbf{K}$ is

[1] Theorem C of Boothby [29] states that for each of the transitive Lie algebras there exists a pair of generating matrices; see Theorem 2.5.

transitive on rays $\{\mathbb{R}_+x\}$. All such maximal compact subgroups are conjugate in $\mathrm{GL}(n,\mathbb{R})$. Any groups $\mathbf{K}', \mathbf{G}'$ conjugate to $\mathbf{K}, \mathbf{G}$ have the same transitivity properties with respect to the corresponding sphere $\{x \in \mathbb{R}^n | \langle x, x\rangle_{\mathbf{K}'} = 1\}$. Without loss of generality assume $\mathbf{K} \subset \mathrm{SO}(n)$. □

The Lie groups $\mathbf{K}$ transitive on spheres, needed by Theorem D.1, were found by Montgomery and Samelson [212] and Borel [33]:

$$\begin{aligned}
&(i) && \mathrm{SO}(n); \\
&(ii) && \mathrm{SU}(k) \subset \mathrm{O}(n) \text{ and } \mathrm{U}(k) \subset \mathrm{O}(n)\ (n = 2k); \\
&(iii)(a) && \mathrm{Sp}(k) \subset \mathrm{O}(n)\ (n = 4k); \\
&(iii)(b) && \mathrm{Sp}(k) \cdot \mathrm{Sp}(1) \subset \mathrm{O}(n)\ (n = 4k); \\
&(iv) && \text{the exceptional groups} \\
& && (a)\ \mathrm{G}_{2(-14)}\ (n = 7), \\
& && (b)\ \mathrm{Spin}(7)\ (n = 8), \\
& && (c)\ \mathrm{Spin}(9)\ (n = 16).
\end{aligned} \tag{D.2}$$

Theorem D.2. *If* $\mathbf{G}$ *is unbounded and has a maximal compact subgroup* $\mathbf{K}$ *whose action as a subgroup of* $\mathrm{Gl}(n,\mathbb{R})$ *is transitive on* $\tilde{S}^{n-1}$*, then* $\mathbf{G}$ *is transitive on* $\mathbb{R}^n_*$*.*

Proof (Sketch). The orbits of $\mathbf{K}$ are spheres, so the orbits of $\mathbf{G}$ are connected subsets of $\mathbb{R}^n$ that are unions of spheres. If there were such spheres with minimal or maximal radii, $\mathbf{G}$ would be bounded. □

A real Lie algebra k is called compact if its Cartan–Killing form (Section B.2.1) is negative definite; the corresponding connected Lie group $\mathbf{K}$ is then compact. To use Theorem D.2 we will need this fact: if the Lie algebra of a real simple group is not a compact Lie algebra then no nontrivial representation of the group is bounded.

The goal is to find connected Lie groups $\tilde{\mathbf{G}}$ that have a representation ρ on $\mathbb{R}^n$ such that $\mathbf{G} = \rho(\tilde{\mathbf{G}})$ is transitive on $\mathbb{R}^n_*$. From transitivity, the linear group action $\mathbf{G} \times \mathbb{R}^n \to \mathbb{R}^n$ is *irreducible*, which implies complete reducibility of the Lie algebra $\mathfrak{g}$ corresponding to the action [143, p. 46]. As a consequence, $\tilde{\mathfrak{g}}$ is reductive, so the only representations needed are faithful ones. Reductivity also implies (using [143, Th. II.11])

$$\mathfrak{g} = \mathfrak{g}_1 + \cdots + \mathfrak{g}_r \oplus \mathfrak{c}, \tag{D.3}$$

where $\mathfrak{c}$ is the center and $\mathfrak{g}_i$ are simple.

The centralizer $\mathfrak{z}$ of an irreducible group representation acting on a real linear space is a division algebra, and any real division algebra is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}_*$. That version of Schur's Lemma is needed because if n is divisible by 2 or 4 the representation may act on $\mathbb{C}^{2m} \simeq \mathbb{R}^n$ or $\boldsymbol{H}^k \simeq \mathbb{R}^{4k}$.

Any center $\mathfrak{c}$ that can occur is a subset of $\mathfrak{z}$ and must be conjugate to one of the following five Lie subalgebras.[2]

$$\begin{gathered}
\mathfrak{c}_0 : \{0\}\,,\ \delta = 0; \\
\mathfrak{c}(n, \mathbb{R}) : \{\lambda I_n |\ \lambda \in \mathbb{R}\}\,,\ \delta = 1; \\
\mathfrak{c}(m, i\mathbb{R}) : \{\boldsymbol{\alpha}(i\lambda I_m) |\ \lambda \in \mathbb{R}\}\,,\ n = 2m,\ \delta = 1; \\
\mathfrak{c}(m, \mathbb{C}) : \left\{\boldsymbol{\alpha}(uI_m + ivI_m) |\ (u, v) \in \mathbb{R}^2\right\},\ n = 2m,\ \delta = 2 \\
\mathfrak{c}(m, (u_0 + iv_0)\mathbb{R}) : \{t\boldsymbol{\alpha}(u_0 I_m + iv_0 I_m) |\ t \in \mathbb{R},\ u_0 \neq 0\}\,,\ n = 2m,\ \delta = 1.
\end{gathered} \tag{D.4}$$

Boothby then points out that the groups in (D.2) are such that in (D.3) we need to consider $r = 1$ or $r = 2$. There are in fact three cases, corresponding to cases I–III in the list of Section D.2:

(I) $r = 1$ and $\mathfrak{g}_1 = \mathfrak{k}_1$, the semisimple part of $\mathfrak{g}$ is simple and compact.

(II) $r = 1$ and the semisimple part of $\mathfrak{g}$ is simple and noncompact.

(III) $r = 2$ and $\mathfrak{k}_1 \oplus \mathfrak{k}_2 = \mathfrak{sp}(p) + \mathfrak{sp}(q)$.

In case I, one needs only to find the correct center $\mathfrak{c}$ from (D.4) in the way just discussed. Case II requires consulting the list of simple groups $\tilde{\mathbf{G}}_1$ having the given $\tilde{\mathbf{K}}$ as maximal compact subgroup [120, Ch. IX, Tables I, II]. Next the compatibility (such as dimensionalities) of the representations of $\tilde{\mathbf{G}}_1$ and $\tilde{\mathbf{K}}$ must be checked (using Tits [277]). For the rest of the story the reader is referred to [29].

Two minor errors in nomenclature in [29] were corrected in both [32, 238]; for **Spin**(5), **Spin**(7) read **Spin**(7), **Spin**(9), respectively.

An additional possibility for a noncompact group of type II was overlooked until recently. Kramer [162, Th. 6.17] shows the transitivity on $\mathbb{R}^{16}_*$ of the noncompact Lie group **Spin**(9, 1) whose Lie algebra $\mathfrak{spin}(9, 1)$ is listed here as II.6. An early proof of the transitivity of $\mathfrak{spin}(9, 1)$ was given by Bryant [45]. It may have been known even earlier.

Definition D.1. A Lie algebra $\mathfrak{g}$ on $\mathbb{R}^n$, n even, is compatible with a complex structure on $\mathbb{R}^n$ if there exists a matrix J such that $J^2 = -I_n$ and $[J, X] = 0$ for each $X \in \mathfrak{g}$; then $[JX, JY] = -[X, Y]$. △

Example D.1. To find a complex structure for $\{A, B\}_{\mathcal{L}}$ with n even (Lie algebras II.2 and II.5 in Section D.2) use the following Mathematica script:

```
Id=IdentityMatrix[n]; J=Array[i,{n,n}]; v=Flatten[J];
Reduce[{J.A-A.J==0,J.B-B.J==0,J.J+Id==0},v];
J=Transpose[Partition[v,n]]
```

The matrices in $\mathfrak{sl}(k, \mathbb{C})$ have zero complex trace: the generators satisfy $\mathrm{tr}(A) = \mathrm{tr}(JA) = 0$ and $\mathrm{tr}(B) = \mathrm{tr}(JB) = 0$.

△

[2] In the fifth type of center, the fixed complex number $u_0 + iv_0$ should have a nonvanishing real part u_0 so that $\Re \exp((u_0 + iv_0)\mathbb{R}) = \mathbb{R}_+$.

D.2 The Transitive Lie Algebras

Here, up to conjugacy, are the Lie algebras $\mathfrak{g}$ transitive on $\mathbb{R}^n_*$, $n \geq 2$, following the lists of Boothby and Wilson [32] and Kramer [162].

Each transitive Lie algebra is listed in the form $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{c}$ where $\mathfrak{g}_0$ is semisimple or 0 and $\mathfrak{c}$ is the center of $\mathfrak{g}$. The list includes some minimal faithful matrix representations; their sources are as noted. The parameter k is any integer greater than 1; representations are $n \times n$; and ℓ is the dimension of $\mathfrak{g}$. The numbering of the Lie algebras in cases I–III is that given in [32], except for the addition of II.6. The dimensions are listed in Table D.1.

n :	2	3	4	5	6	7	8	9	16
I.1	2	4	7	11	16	22	29	37	121
I.2			4, 5		9, 10		16, 17		65
I.3							11, 12		37, 38
I.4						14		40	
I.5							22		
I.6									37
I.7						15			
II.1	3, 4	8, 9	15, 16	24, 25	35, 36	48, 49	63, 64	80, 81	255, 256
II.2			6, 7, 8		16, 17, 18		30, 31, 32		126, 127, 128
II.3							15, 16, 17		63, 64, 65
II.4			10, 11		21, 22		36, 37		136, 137
II.5							20, 21, 22		72, 73, 74
II.6									45
III							18, 19		66, 67

Table D.1. Dimension ℓ of each Lie algebra transitive on $\mathbb{R}^n_*$ for $1 < n \leq 9$ and $n = 16$. If n is odd and $n \neq 7$ the only transitive Lie groups are $GL^+(n, \mathbb{R})$, $SL(n, \mathbb{R})$, and $SO(n) \times \mathfrak{c}(n, \mathbb{R})$, which are easy to recognize.

I. **Semisimple ideal is simple and compact**[3]

(1) $\mathfrak{g} = \mathfrak{so}(k) \oplus \mathfrak{c} \subset \mathfrak{gl}(k, \mathbb{R}),\ n = k;\ \mathfrak{c} = \mathfrak{c}(k, \mathbb{R});\ \ell = 1 + k(k-1)/2;$

$\mathfrak{so}(k) = \{X \in \mathfrak{gl}(k, \mathbb{R}) |\, X^\top + X = 0\}.$

(2) $\mathfrak{g} = \mathfrak{su}(k) \oplus \mathfrak{c} \subset \mathfrak{gl}(2k, \mathbb{R}),\ n = 2k;$

$\mathfrak{c} = \mathfrak{c}(k, \mathbb{C})$ or $\mathfrak{c}(k, (u_0 + iv_0)\mathbb{R}),\ \ell = k^2 - 1 + \delta(\mathfrak{c});$

$\mathfrak{su}(k) = \{\boldsymbol{\alpha}(A + iB) |\ A^\top = -A, B^\top = B, \operatorname{tr}(B) = 0\}.$

(3) $\mathfrak{g} = \mathfrak{sp}(k) \oplus \mathfrak{c} \subset \mathfrak{gl}(4k, \mathbb{R}),\ n = 4k,$

$\mathfrak{c} = \mathfrak{c}(2k, \mathbb{C})$ or $\mathfrak{c}(2k, (u_0 + iv_0)\mathbb{R}),\ \ell = 2k^2 + k + \delta(\mathfrak{c}).$

$$\mathfrak{sp}(k) = \left\{ \boldsymbol{\beta}(A + iB + jC + kD) \;\middle|\; \begin{matrix} A^\top = -A, B^\top = B, \\ C^\top = C, D^\top = D \end{matrix} \right\}.$$

[3] In part I, $\mathfrak{g}$ is simple with the exceptions I.1 for $n = 3$ and I.4.

(4) $\mathfrak{g} = \mathfrak{sp}(k) \oplus \mathbb{H}$, $\mathfrak{sp}(k)$ as in (I.3), $n = 4k$, $\ell = 2k^2 + k + 4$.

(5) $\mathfrak{g} = \mathfrak{spin}(7) \oplus \mathfrak{c} \subset \mathfrak{gl}(8, \mathbb{R})$, $\mathfrak{c} = \mathfrak{c}(8, \mathbb{R})$, $n = 8, \ell = 22$.

Kimura [159, Sec. 2.4] has $\mathfrak{spin}(7) =$

$$\left\{\begin{bmatrix} U_1 & -u_{16} & -u_{17} & -u_{18} & 0 & u_{21} & -u_{20} & u_{19} \\ u_{10} & U_2 & u_8 & u_{14} & -u_{21} & 0 & -u_{12} & u_7 \\ u_{11} & u_6 & U_3 & u_{15} & u_{20} & u_{12} & 0 & -u_4 \\ u_{12} & -u_{20} & -u_{21} & U_4 & -u_{19} & -u_7 & u_4 & 0 \\ 0 & u_{15} & -u_{14} & u_{13} & -U_1 & -u_{10} & -u_{11} & -u_{12} \\ -u_{15} & 0 & -u_{18} & -u_3 & u_{16} & -U_2 & -u_6 & u_{20} \\ u_{14} & u_{18} & 0 & u_2 & u_{17} & -u_8 & -U_3 & u_{21} \\ -u_{13} & u_3 & -u_2 & 0 & u_{18} & -u_{14} & -u_{15} & -U_4 \end{bmatrix}\right\},$$

$$\text{where } U_1 = \frac{-u_1 - u_5 - u_9}{2}, U_2 = \frac{u_1 + u_5 - u_9}{2},$$
$$U_3 = \frac{u_1 - u_5 + u_9}{2}, U_4 = \frac{u_1 - u_5 - u_9}{2}.$$

Using (D.1) a basis for $\mathfrak{spin}(7)$ can be generated by the six 8×8 matrices $I \otimes J \otimes L$, $L \otimes J \otimes K$, $I \otimes K \otimes J$, $J \otimes L \otimes I$, $J \otimes K \otimes I$, and $J \otimes K \otimes K$.

(6) $\mathfrak{g} = \mathfrak{spin}(9) \oplus \mathfrak{c} \subset \mathfrak{gl}(16, \mathbb{R})$, $\mathfrak{c} = \mathfrak{c}(16, \mathbb{R})$, $n = 16$, $\ell = 37$.

A 36-matrix basis for $\mathfrak{spin}(9)$ is generated (my computation) by the six 16×16 matrices

$J \otimes J \otimes J \otimes I$, $I \otimes L \otimes J \otimes L$, $J \otimes L \otimes L \otimes L$,
$I \otimes L \otimes L \otimes J$, $J \otimes L \otimes I \otimes I$, $J \otimes K \otimes I \otimes I$;

or see Kimura [159, Example 2.19].

(7) $\mathfrak{g} = \mathfrak{g}_{2(-14)} \oplus \mathfrak{c} \subset \mathfrak{gl}(7, \mathbb{R})$, $\mathfrak{c} = \mathfrak{c}(7, \mathbb{R})$; $n = 7$, $\ell = 15$; $\mathfrak{g}_{2(-14)} =$

$$\left\{\begin{bmatrix} 0 & -u_1 & -u_2 & -u_3 & -u_4 & -u_5 & -u_6 \\ u_1 & 0 & -u_7 & -u_8 & -u_9 & -u_{10} & -u_{11} \\ u_2 & u_7 & 0 & -u_5 - u_9 & -u_6 - u_8 & -u_{11} + u_3 & u_4 + u_{10} \\ u_3 & u_8 & u_5 - u_9 & 0 & -u_{12} & -u_{13} & -u_{14} \\ u_4 & u_9 & u_6 + u_8 & u_{12} & 0 & -u_1 - u_{14} & -u_2 + u_{13} \\ u_5 & u_{10} & u_{11} - u_3 & u_{13} & u_1 + u_{14} & 0 & -u_7 - u_{12} \\ u_6 & u_{11} & -u_4 - u_{10} & u_{14} & u_2 - u_{13} & u_7 + u_{12} & 0 \end{bmatrix}\right\}.$$

$\mathbf{G}_2$ is an "exceptional" Lie group; for its Lie algebra see Gross [111], Baez [16]; $\mathfrak{g}_{2(-14)}$ is so called because -14 is the signature of its Killing form $\chi(X, Y)$.

II. **Semisimple ideal is simple and noncompact**

(1) $\mathfrak{g} = \mathfrak{sl}(k, \mathbb{R}) \oplus \mathfrak{c} \subset \mathfrak{gl}k, \mathbb{R})$, $n = k$;
$\mathfrak{c} = \mathfrak{c}_0$ or $\mathfrak{c}(k, \mathbb{R})$; $\ell = k^2 - 1 + \delta(\mathfrak{c})$;
$\mathfrak{sl}(k, \mathbb{R}) = \{X \subset \mathfrak{gl}(k, \mathbb{R}) | \operatorname{tr}(X) = 0\}$;

(2) $\mathfrak{g} = \mathfrak{sl}(k, \mathbb{C}) \oplus \mathfrak{c} \subset \mathfrak{gl}(2k, \mathbb{R}),\ n = 2k,$

$\mathfrak{c} = \mathfrak{c}_0, \quad \mathfrak{c}(n, \mathbb{R}),\ \mathfrak{c}(k, i\mathbb{R}),\ \mathfrak{c}(m, \mathbb{C}),$ or $\mathfrak{c}(k, (u_0 + iv_0)\mathbb{R});$

$\ell = 2(k^2 - 1) + \delta(\mathfrak{c});$

$\mathfrak{sl}(k, \mathbb{C}) = \{\boldsymbol{\alpha}(A + iB) \mid A, B \in \mathbb{R}^{k\times k}, \operatorname{tr}(A) = 0 = \operatorname{tr}(B)\}.$

(3) $\mathfrak{g} = \mathfrak{sl}(k, \mathbb{H}) \oplus \mathfrak{c} \subset \mathfrak{gl}(4k, \mathbb{R}),\ n = 4k,$

$\mathfrak{c} = \mathfrak{c}_0, \mathfrak{c}(4k, \mathbb{R}), \mathfrak{c}(2k, i\mathbb{R}), \mathfrak{c}(2k, \mathbb{C}),$ or $\mathfrak{c}(2k, (u_0 + iv_0)\mathbb{R}),$

$\ell = 4k^2 - 1 + \delta(\mathfrak{c});\ \mathfrak{sl}(k, \mathbb{H}) = \{\boldsymbol{\alpha}(\mathfrak{su}^*(2k))\},$ where

$$\mathfrak{su}^*(2k) = \left\{ \begin{bmatrix} A & B \\ -\bar{B} & \bar{A} \end{bmatrix} \,\middle|\, \begin{matrix} A, B \in \mathbb{C}^{k\times k} \\ \operatorname{tr} A + \operatorname{tr} \bar{A} = 0 \end{matrix} \right\}.$$

(4) $\mathfrak{g} = \mathfrak{sp}(k, \mathbb{R}) \oplus \mathfrak{c} \subset \mathfrak{gl}(2k, \mathbb{R}),\ n = 2k,\ \mathfrak{c} = \mathfrak{c}_0$ or $\mathfrak{c}(n, \mathbb{R});$

$$\ell = 2k^2 + k + \delta(\mathfrak{c});\ \mathfrak{sp}(k, \mathbb{R}) = \left(\begin{bmatrix} A & B \\ C & -A^\top \end{bmatrix} \,\middle|\, \begin{matrix} A, B, C \in \mathbb{R}^{k\times k} \\ B \text{ and } C \text{ symmetric} \end{matrix} \right\}$$

is a real symplectic algebra.

(5) $\mathfrak{g} = \mathfrak{sp}(k, \mathbb{C}) \oplus \mathfrak{c} \subset \mathfrak{gl}(4k, \mathbb{R}),\ n = 4k,$

$\mathfrak{c} = \mathfrak{c}_0, \mathfrak{c}(4k, \mathbb{R}), \mathfrak{c}(2k, i\mathbb{R}), \mathfrak{c}(2k, \mathbb{C}),$ or $\mathfrak{c}(2k, (u_0 + iv_0)\mathbb{R});$

$\ell = 4k^2 + 2k + \delta(\mathfrak{c});$

$$\mathfrak{sp}(k, \mathbb{C}) = \left\{ \boldsymbol{\alpha}\left(\begin{bmatrix} A & B \\ C & -A^\top \end{bmatrix} + i \begin{bmatrix} D & E \\ F & -D \end{bmatrix} \right) \,\middle|\, \begin{matrix} A, B, C, D, E, F \in \mathbb{R}^{k\times k} \\ B, C, E, F, \text{ symmetric} \end{matrix} \right\}.$$

(6) $\mathfrak{g} = \mathfrak{spin}(9, 1) \oplus \mathfrak{c} \subset \mathfrak{gl}(16, \mathbb{R});\ \ n = 16,$

$\mathfrak{c} = \mathfrak{c}_0$ or $\mathfrak{c}(16, \mathbb{R}), \ell = 45 + \delta(\mathfrak{c});$ see Kramer [162].

$$\mathfrak{spin}(9, 1) = \left\{ \begin{bmatrix} \rho_1'(a) + \lambda I_8 & R_{\mathbf{w}}^\dagger \\ L_{\mathbf{x}}^\dagger & \rho_3'(a) - \lambda I_8 \end{bmatrix} \,\middle|\, \begin{matrix} \lambda \in \mathbb{R}, a \in \mathfrak{spin}(8), \\ \mathbf{x}, \mathbf{w} \in \mathbb{O} \end{matrix} \right\}$$

in the representation explained by Bryant [45], who also points out its transitivity on $\mathbb{R}^{16}_*$. Here $\mathbb{O}$ denotes the octonions, a normed 8-dimensional nonassociative division algebra; the multiplication table for $\mathbb{O}$ can be seen as a matrix representation on $\mathbb{R}^8$ of left (right) multiplication $L_{\mathbf{x}}$, $[R_{\mathbf{w}}]$ by $\mathbf{x}, \mathbf{w} \in \mathbb{O}$; their octonion-conjugates $L_{\mathbf{x}}^\dagger$, $[R_{\mathbf{w}}^\dagger]$ are used in the $\mathfrak{spin}(9, 1)$ representation. To see that $\ell = 45$, note that $\dim \mathfrak{spin}(8) = 28$, $\dim \mathbb{O} = 8$, and $\dim \mathbb{R} = 1$. The Lie algebra isomorphisms $\rho_i' : \mathfrak{spin}(8) \to \mathfrak{so}(8)$ are defined in [45]. The maximal compact subgroup in $\mathfrak{spin}(9, 1)$ is $\mathfrak{spin}(9)$, see I.6 above. Baez [16] gives the isomorphism $\mathfrak{spin}(9, 1) \cong \mathfrak{sl}(\mathbb{O}, 2)$; here $\mathbb{R}^{16} \cong \mathbb{O} \oplus \mathbb{O}$, and the octonion multiplication table provides the representation.

III. **The remaining Lie algebras**

From Boothby and Wilson [32]:

$$\mathfrak{g} = \mathfrak{sl}(k, \mathbb{H}) \oplus \mathfrak{su}(2) \oplus \mathfrak{c},\ n = 4k;\ \mathfrak{c} = \mathfrak{c}_0 \text{ or } \mathfrak{c}(n, \mathbb{R});\ \ell = 4k^2 + 2 + \delta(\mathfrak{c}).$$

The Lie algebra $\mathfrak{su}(2)$ can be identified with $\mathfrak{sp}(1)$ to help comparison with the list in Kramer [162]. (For more about $\mathfrak{sp}(k)$ see Chevalley [58, Ch. 1].)

"The representative $\mathfrak{sl}(k, \mathbb{H}) \oplus \mathfrak{su}(2)$ is to be regarded as the subalgebra of $\mathfrak{gl}(4k, \mathbb{R})$ obtained by the natural left action of $\mathfrak{sl}(k)$ on $\mathbb{H}^k \cong \mathbb{R}^{4k}$ and the right action of $\mathfrak{su}(2)$ as a multiplication by pure quaternions on $\mathbb{H}^k$. " [32]

Problem D.1. The algorithm in Boothby and Wilson [32] should be realizable with a modern symbolic computational system.

D.2.1 The Isotropy Subalgebras

It would be interesting and useful to identify all the isotropy subalgebras $\mathfrak{iso}(\mathfrak{g})$ for the transitive Lie algebras. Here are a few, numbered in correspondence with the above list. For II.1, compare the group in Example 2.18 and the definition in (3.30). Let

$$\mathfrak{aff}(n-1, \mathbb{R})_R := \left\{ X = \begin{bmatrix} A & 0 \\ b^\top & 0 \end{bmatrix} \middle| \; A \in \mathfrak{gl}(n-1, \mathbb{R}), \; b \in \mathbb{R}^{n-1} \right\}.$$

The matrices in $\mathfrak{aff}(n-1, \mathbb{R})_R$ are the transposes of the matrices in $\mathfrak{aff}(n-1, \mathbb{R})$; their nullspace is the span of $\delta_n := \mathrm{col}(0, 0, \ldots, 1)$.

I.1: $\mathfrak{iso}(\mathfrak{so}(k) \oplus \mathfrak{c}(k, \mathbb{R}) = \mathfrak{so}(k-1)$.

I.5: $\mathfrak{iso}(\mathfrak{spin}(7)) = \mathfrak{g}_2$.

I.6: $\mathfrak{iso}(\mathfrak{spin}(9) \oplus \mathfrak{c}(16, \mathbb{R})) = \mathfrak{spin}(7)$.

II.1: If $\mathfrak{c} = \mathfrak{c}(k, \mathbb{R}), \; \mathfrak{iso}(\mathfrak{gl}(n, \mathbb{R})) = \mathfrak{aff}(n-1, \mathbb{R})_R$.

II.1: If $\mathfrak{c} = \mathfrak{c}_0, \; \mathfrak{iso}(\mathfrak{sl}(n, \mathbb{R})) = \{X \,|\, X \in \mathfrak{aff}(n-1, \mathbb{R})_R, \; \mathrm{tr}(X) = 0\}$.

References

1. A. A. Agrachev and D. Liberzon, "Lie-algebraic stability criteria for switched systems," *SIAM J. Control Opt.*, vol. 40, pp. 253–259, 2001.
2. F. Albertini and D. D'Alessandro, "The Lie algebra structure of spin systems and their controllability properties," in *Proc. 40th IEEE Conf. Decis. Contr.* New York: IEEE Pubs., 2001, pp. 2599–2604.
3. F. Albertini and D. D'. Alessandro, "Notions of controllability for quantum mechanical systems," in *Proc. 40th IEEE Conf. Decis. Contr.* New York: IEEE Pubs., 2001, pp. 1589–1594.
4. C. Altafini, "Controllability of quantum mechanical systems by root space decomposition of $\mathfrak{su}(n)$," *J. Math. Phys.*, vol. 43, no. 5, pp. 2051–2062, 2002. [Online]. Available: http://link.aip.org/link/?JMP/43/2051/1
5. C. Altafini, "The reachable set of a linear endogenous switching system," *Syst. Control Lett.*, vol. 47, no. 4, pp. 343–353, 2002.
6. C. Altafini, "Explicit Wei-Norman formulae for matrix Lie groups via Putzer's method," *Syst. Control Lett.*, vol. 54, no. 11, pp. 1121–1130, 2006. [Online]. Available: http://www.sissa.it/~altafini/papers/wei-norman.pdf
7. P. Antsaklis and A. N. Michel, *Linear Systems*. New York: McGraw-Hill, 1997.
8. M. A. Arbib and E. G. Manes, "Foundations of system theory: The Hankel matrix." *J. Comp. Sys. Sci.*, vol. 20, pp. 330–378, 1980.
9. L. Arnold and W. Kliemann, "Qualitative theory of stochastic systems," in *Probabilistic Analysis and Related Topics, Vol. 3*, A. T. Bharucha-Reid, Ed. New York: Academic Press, 1983, pp. 1–79.
10. V. Arnol'd, *Ordinary Differential Equations*. New York: Springer-Verlag, 1992.
11. V. Ayala and L. A. B. San Martín, "Controllability of two-dimensional bilinear systems: restricted controls and discrete-time," *Proyecciones*, vol. 18, no. 2, pp. 207–223, 1999.
12. V. Ayala and J. Tirao, "Linear systems on Lie groups and controllability," in *Differential Geometry and Control*, G. Ferreyra, R. Gardner, H. Hermes, and H. Sussmann, Eds. Providence: AMS, 1999, pp. 47–64.
13. A. Bacciotti, "On the positive orthant controllability of two-dimensional bilinear systems," *Syst. Control Lett.*, vol. 3, pp. 53–55, 1983.
14. A. Bacciotti and P. Boieri, "A characterization of single-input planar bilinear systems which admit a smooth stabilizer," *Syst. Control Lett.*, vol. 16, 1991.
15. A. Bacciotti and P. Boieri, "Stabilizability of oscillatory systems: a classical approach supported by symbolic computation," *Matematiche (Catania)*, vol. 45, no. 2, pp. 319–335 (1991), 1990.

D.L. Elliott, *Bilinear Control Systems,* Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_13,

16. J. Baez, "The octonions," *Bull. Amer. Math. Soc.*, vol. 39, no. 2, pp. 145–205, 2001, errata: *Bull. Amer. Math. Soc.*, vol. 42, p. 213 (2005).
17. J. Baillieul, "Geometric methods for nonlinear optimal control problems." *J. Optim. Theory Appl.*, vol. 25, pp. 519–548, 1978.
18. J. A. Ball, G. Groenewald, and T. Malakorn, "Structured noncommutative multidimensional linear systems," *SIAM J. Control Optim.*, vol. 44, no. 4, pp. 1474–1528 (electronic), 2005.
19. J. M. Ball and M. Slemrod, "Feedback stabilization of distributed semilinear control systems." *Appl. Math. Optim.*, vol. 5, no. 2, pp. 169–179, 1979.
20. H. T. Banks, R. P. Miech, and D. J. Zinberg, "Nonlinear systems in models for enzyme cascades," in *Variable Structure Systems with Applications to Economics and Biology. (Proceedings 2nd U.S. – Italy Seminar, Portland, Oregan*, A. Ruberti and R. R. Mohler, Eds., Berlin: Springer-Verlag, 1974, pp. 265–277.
21. E. A. Barbashin [Barbašin], *Introduction to the Theory of Stability*, T. Lukes, Ed. Groningen: Wolters-Noordhoff Publishing, 1970, translation of *Vvedenie v teoriyu ustoichivosti*.
22. E. A. Barbašin, *Vvedenie v teoriyu ustoichivosti [Introduction to the Theory of Stability].* Izdat. "Nauka", Moscow, 1967.
23. C. J. B. Barros, J. R. Gonçalves Filho, O. G. do Rocio, and L. A. B. San Martin, "Controllability of two-dimensional bilinear systems," *Proyecciones Revista de Matematica*, vol. 15, 1996.
24. R. H. Bartels and G. W. Stewart, "Solution of the equation AX + XB =C," *Comm. Assoc. Comp. Mach.*, vol. 15, pp. 820–826, 1972.
25. R. E. Bellman, *Introduction to Matrix Analysis.* New York: McGraw-Hill, (Reprint of 1965 edition) 1995.
26. V. D. Blondel and A. Megretski, Eds., *Unsolved Problems in Mathematical Systems & Control Theory*. Princeton, NJ: Princeton University Press, 2004. [Online]. Available: http://www.inma.ucl.ac.be/~blondel/books/openprobs/
27. J. Bochnak, M. Coste, and M.-F. Roy, *Géométrie Algébrique réelle*. Berlin: Springer-Verlag, 1987.
28. B. Bonnard, V. Jurdjevic, I. Kupka, and G. Sallet, "Transitivity of families of invariant vector fields on the semidirect products of Lie groups," *Trans. Amer. Math. Soc.*, vol. 271, no. 2, pp. 525–535, 1982.
29. W. M. Boothby, "A transitivity problem from control theory." *J. Differ. Equ.*, vol. 17, no. 296–307, 1975.
30. W. M. Boothby, "Some comments on positive orthant controllability of bilinear systems." *SIAM J. Control Optim.*, vol. 20, pp. 634–644, 1982.
31. W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, 2nd ed. New York: Academic Press, 2003.
32. W. M. Boothby and E. N. Wilson, "Determination of the transitivity of bilinear systems," *SIAM J. Control Optim.*, vol. 17, pp. 212–221, 1979.
33. A. Borel, "Le plan projectif des octaves et les sphères comme espaces homogènes," *C. R. Acad. Sci. Paris*, vol. 230, pp. 1378–1380, 1950.
34. R. W. Brockett, "Differential geometric methods in control," Division of Engineering and Applied Physics, Harvard University, Tech. Rep. 628, 1971.
35. R. W. Brockett, "On the algebraic structure of bilinear systems." in *Theory and Applications of Variable Structure Systems*, R. Mohler and A. Ruberti, Eds., New York: Academic Press, 1972, pp. 153–168.
36. R. W. Brockett, "System theory on group manifolds and coset spaces." *SIAM J. Control Optim.*, vol. 10, pp. 265–284, 1972.
37. R. W. Brockett, "Lie theory and control systems defined on spheres." *SIAM J. Appl. Math.*, vol. 25, pp. 213–225, 1973.
38. R. W. Brockett, "Volterra series and geometric control theory." *Automatica—J. IFAC*, vol. 12, pp. 167–176, 1976.

39. R. W. Brockett, "On the reachable set for bilinear systems," in *Variable Structure Systems with Applications to Economics and Biology. (Proceedings 2nd U.S. – Italy Seminar, Portland, Oregan*, A. Ruberti and R. R. Mohler, Eds., Berlin: Springer-Verlag, 1975, pp. 54–63.
40. H. W. Broer, "KAM theory: the legacy of A. N. Kolmogorov's 1954 paper," *Bull. Amer. Math. Soc. (N.S.)*, vol. 41, no. 4, pp. 507–521 (electronic), 2004.
41. W. C. Brown, *Matrices over Commutative Rings*. New York: Marcel Dekker Inc., 1993.
42. C. Bruni, G. D. Pillo, and G. Koch, "On the mathematical models of bilinear systems," *Ricerche Automat.*, vol. 2, no. 1, pp. 11–26, 1971.
43. C. Bruni, "Bilinear systems: an appealing class of nearly linear systems in theory and application." *IEEE Trans. Automat. Control*, vol. AC-19, pp. 334–348, 1974.
44. A. D. Brunner and G. D. Hess, "Potential problems in estimating bilinear time-series models," *J. Econ. Dyn. Control*, vol. 19, pp. 663–681, 1995. [Online]. Available: http://ideas.repec.org/a/eee/dyncon/v19y1995i4p663-681.html
45. R. L. Bryant. Remarks on spinors in low dimension. [Online]. Available: http://www.math.duke.edu/~bryant/Spinors.pdf
46. V. I. Buyakas, "Optimal control of systems with variable structure." *Automat. Remote Control*, vol. 27, pp. 579–589, 1966.
47. G. Cairns and É. Ghys, "The local linearization problem for smooth SL(n)-actions," *Enseign. Math. (2)*, vol. 43, no. 1–2, pp. 133–171, 1997.
48. F. Cardetti and D. Mittenhuber, "Local controllability for linear systems on Lie groups," *J. Dyn. Control Syst.*, vol. 11, no. 3, pp. 353–373, 2005.
49. S. Čelikovský, "On the stabilization of the homogeneous bilinear systems," *Syst. Control Lett.*, vol. 21, pp. 503–510, 1993.
50. S. Čelikovský, "On the global linearization of nonhomogeneous bilinear systems," *Syst. Control Lett.*, vol. 18, no. 18, pp. 397–402, 1992.
51. S. Čelikovský and A. Vaněček, "Bilinear systems and chaos," *Kybernetika (Prague)*, vol. 30, pp. 403–424, 1994.
52. R. Chabour, G. Sallet, and J. C. Vivalda, "Stabilization of nonlinear two-dimensional systems: A bilinear approach." *Math. Control Signals Systems*, vol. 6, no. 3, pp. 224–246, 1993.
53. K.-T. Chen, "Integrations of paths, geometric invariants and a generalized Baker-Hausdorff formula," *Ann. Math.*, vol. 65, pp. 163–178, 1957.
54. K.-T. Chen, "Equivalence and decomposition of vector fields about an elementary critical point," *Amer. J. Math.*, vol. 85, pp. 693–722, 1963.
55. K.-T. Chen, "Algebras of iterated path integrals and fundamental groups," *Trans. Amer. Math. Soc.*, vol. 156, pp. 359–379, 1971.
56. G.-S. J. Cheng, "Controllability of discrete and continuous-time bilinear systems," Sc.D. Dissertation, Washington University, St. Louis, Missouri, December 1974.
57. G.-S. J. Cheng, T.-J. Tarn, and D. L. Elliott, "Controllability of bilinear systems," in *Variable Structure Systems with Applications to Economics and Biology. (Proceedings 2nd U.S. – Italy Seminar, Portland, Oregan*, A. Ruberti and R. R. Mohler, Eds., Berlin: Springer-Verlag, 1975, pp. 83–100.
58. C. Chevalley, *Theory of Lie Groups. I*. Princeton, NJ: Princeton University Press, 1946.
59. W. L. Chow, "Über systeme von linearen partiellen differential-gleichungen erster ordnung," *Math. Annalen*, vol. 117, pp. 98–105, 1939.
60. J. M. C. Clark, "An introduction to stochastic differential equations on manifolds," in *Geometric Methods in System Theory*. Dordrecht, The Netherlands: D. Reidel, 1973, pp. 131–149.

61. J. W. Clark, "Control of quantum many-body dynamics: Designing quantum scissors," in *Condensed Matter Theories, Vol. 11*, E. V. Ludeña, P. Vashishta, and R. F. Bishop, Eds. Commack, NY: Nova Science Publishers, 1996, pp. 3–19.
62. F. H. Clarke, Y. S. Ledyaev, and R. J. Stern, "Asymptotic stability and smooth Lyapunov functions," *J. Differ. Equ.*, vol. 149, no. 1, pp. 69–114, 1998.
63. F. Colonius and W. Kliemann, *The Dynamics of Control*. Boston: Birkhäuser, 1999.
64. L. Conlon, *Differentiable Manifolds: A First Course*. Boston, MA: Birkhäuser, 1993.
65. M. J. Corless and A. E. Frazho, *Linear Systems and Control: An Operator Perspective*. New York: Marcel Dekker, 2003.
66. J. Cortés, "Discontinuous dynamical systems: A tutorial on solutions, nonsmooth analysis and stability," *IEEE Contr. Syst. Magazine*, vol. 28, no. 3, pp. 36–73, June 2008.
67. D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*, 2nd ed. New York: Springer-Verlag, 1997.
68. P. d'Alessandro, "Bilinearity and sensitivity in macroeconomy," in *Variable Structure Systems with Applications to Economics and Biology. (Proceedings 2nd U.S. – Italy Seminar, Portland, Oregan*, A. Ruberti and R. R. Mohler, Eds., Berlin: Springer-Verlag, 1975.
69. P. d'Alesandro, A. Isidori, and A. Ruberti, "Realization and structure theory of bilinear systems." *SIAM J. Control Optim.*, vol. 12, pp. 517–535, 1974.
70. D. D'Alessandro, "Small time controllability of systems on compact Lie groups and spin angular momentum," *J. Math. Phys.*, vol. 42, pp. 4488–4496, 2001.
71. D. D'Alessandro, *Introduction to Quantum Control and Dynamics*. Boca Raton: Taylor & Francis, 2007.
72. D. D'Alessandro and M. Dahleh, "Optimal control of two-level quantum systems," *IEEE Trans. Automat. Control*, vol. 46, pp. 866–876, 2001.
73. G. Dirr and H. K. Wimmer, "An Eneström-Kakeya theorem for Hermitian polynomial matrices," *IEEE Trans. Automat. Control*, vol. 52, no. 11, pp. 2151–2153, November 2007.
74. J. D. Dollard and C. N. Friedman, *Product Integration with Applications to Differential Equations*. Reading: Addison-Wesley, 1979.
75. O. G. do Rocio, L. A. B. San Martin, and A. J. Santana, "Invariant cones and convex sets for bilinear control systems and parabolic type of semigroups," *J. Dyn. Control Syst.*, vol. 12, no. 3, pp. 419–432, 2006.
76. G. Dufrénot and V. Mignon, *Recent Developments in Nonlinear Cointegration with Applications to Macroeconomics and Finance*. Berlin: Springer-Verlag, 2002.
77. J. J. Duistermaat and J. A. Kolk, *Lie Groups*. Berlin: Springer-Verlag, 2000.
78. E. B. Dynkin, *Normed Lie Algebras and Analytic Groups*. Providence, RI: American Mathematical Society, 1953.
79. M. Eisen, *Mathematical Models in Cell Biology and Cancer Chemotherapy*. Springer-Verlag: Lec. Notes in Biomath. **30**, 1979.
80. D. L. Elliott, "A consequence of controllability." *J. Differ. Equ.*, vol. 10, pp. 364–370, 1971.
81. D. L. Elliott, "Controllable nonlinear systems driven by white noise," Ph.D. Dissertation, University of California at Los Angeles, 1969. [Online]. Available: https://drum.umd.edu/dspace/handle/1903/6444
82. D. L. Elliott, "Diffusions on manifolds, arising from controllable systems," in *Geometric Methods in System Theory*. Dordrecht, The Netherlands: D. Reidel, 1973, pp. 285–294.
83. D. L. Elliott, "Bilinear systems," in *Wiley Encyclopedia of Electrical Engineering*, J. Webster, Ed. New York: Wiley, 1999, vol. 2, pp. 308–323.
84. D. L. Elliott, "A controllability counterexample," *IEEE Trans. Automat. Control*, vol. 50, no. 6, pp. 840–841, June 2005.

85. D. L. Elliott and T.-J. Tarn, "Controllability and observability for bilinear systems," SIAM Natl. Meeting, Seattle, Washington, Dec. 1971; Report CSSE-722, Dept. of Systems Sci. and Math., Washington University, St. Louis, Missouri, 1971.
86. G. Escobar, R. Ortega, H. Sira-Ramirez, J.-P. Vilain, and I. Zein, "An experimental comparison of several nonlinear controllers for power converters," *IEEE Contr. Syst. Magazine*, vol. 19, no. 1, pp. 66–82, February 1999.
87. M. E. Evans and D. N. P. Murthy, "Controllability of a class of discrete time bilinear systems," *IEEE Trans. Automat. Control*, vol. 22, pp. 78–83, February 1977.
88. W. Favoreel, B. De Moor, and P. Van Overschee, "Subspace identification of bilinear systems subject to white inputs," *IEEE Trans. Automat. Control*, vol. 44, no. 6, pp. 1157–1165, 1999.
89. G. Ferreyra, R. Gardner, H. Hermes, and H. Sussmann, Eds., *Differential Geometry and Control*, (Proc. Symp. Pure Math. Vol. 64). Providence: Amer. Math. Soc., 1999.
90. A. F. Filippov, "Differential equations with discontinuous right-hand side," in *Amer. Math. Soc. Translations*, 1964, vol. 42, pp. 199–231.
91. M. Flato, G. Pinczon, and J. Simon, "Nonlinear representations of Lie groups," *Ann. Scient. Éc. Norm. Sup. (4th ser.)*, vol. 10, pp. 405–418, 1977.
92. M. Fliess, "Un outil algébrique: les séries formelles non commutatives." in *Mathematical Systems Theory*, G. Marchesini and S. K. Mitter, Eds. Berlin: Springer-Verlag, 1975, pp. 122–148.
93. M. Fliess, "Fonctionelles causales non linéaires et indéterminés noncommutatives." *Bull. Soc. Math. de France*, vol. 109, no. 1, pp. 3–40, 1981. [Online]. Available: http://www.numdam.org./item?id=BSMF_1981__109__3_0
94. E. Fornasini and G. Marchesini, "Algebraic realization theory of bilinear discrete-time input-output maps," *J. Franklin Inst.*, vol. 301, no. 1–2, pp. 143–159, 1976, recent trends in systems theory.
95. E. Fornasini and G. Marchesini, "A formal power series approach to canonical realization of bilinear input-output maps," in *Mathematical Systems Theory*, G. Marchesini and S. K. Mitter, Eds. Berlin: Springer_Verlag, 1976, pp. 149–164.
96. M. Frayman, "Quadratic differential systems; a study in nonlinear systems theory," Ph.D. dissertation, University of Maryland, College Park, Maryland, June 1974.
97. M. Frayman, "On the relationship between bilinear and quadratic systems," *IEEE Trans. Automat. Control*, vol. 20, no. 4, pp. 567–568, 1975.
98. A. E. Frazho, "A shift operator approach to bilinear system theory." *SIAM J. Control Optim.*, vol. 18, no. 6, pp. 640–658, 1980.
99. B. A. Frelek and D. L. Elliott, "Optimal observation for variable–structure systems," in *Proc. VI IFAC Congress, Boston, Mass.*, vol. I, 1975, paper 29.5.
100. S. Friedland, "Simultaneous similarity of matrices," *Adv. Math.*, vol. 50, pp. 189–265, 1983.
101. F. R. Gantmacher, *The Theory of Matrices*. AMS Chelsea Pub. [Reprint of 1977 ed.], 2000, vol. I and II.
102. J. Gauthier and G. Bornard, "Controlabilité des systèmes bilinéaires," *SIAM J. Control Opt.*, vol. 20, pp. 377–384, 1982.
103. J.-P. Gauthier and I. Kupka, "A separation principle for bilinear systems with dissipative drift," *IEEE Trans. Automat. Control*, vol. 37, no. 12, pp. 1970–1974, 1992.
104. S. Gibson, A. Wills, and B. Ninness, "Maximum-likelihood parameter estimation of bilinear systems," *IEEE Trans. Automat. Control*, vol. 50, no. 10, pp. 1581–1596, October 2005.
105. T. Goka, T.-J. Tarn, and J. Zaborszky, "Controllability of a class of discrete-time bilinear systems." *Automatica—J. IFAC*, vol. 9, pp. 615–622, 1973.
106. W. M. Goldman, "Re: Stronger version," personal communication, March 16 2004.
107. M. Goto, "On an arcwise connected subgroup of a Lie group," *Proc. Amer. Math. Soc.*, vol. 20, pp. 157–162, 1969.

108. O. M. Grasselli and A. Isidori, "Deterministic state reconstruction and reachability of bilinear control processes." in *Proc. Joint Automatic Control Conf., San Francisco, June 22-25, 1977*. New York: IEEE Pubs., June 22–25 1977, pp. 1423–1427.
109. O. M. Grasselli and A. Isidori, "An existence theorem for observers of bilinear systems." *IEEE Trans. Automat. Control*, vol. AC-26, pp. 1299–1300, 1981, erratum, same Transactions AC-27:284 (February 1982).
110. G.-M. Greuel, G. Pfister, and H. Schönemann, SINGULAR *2.0. A Computer Algebra System for Polynomial Computations*, Centre for Computer Algebra, University of Kaiserslautern, 2001. [Online]. Available: http://www.singular.uni-kl.de
111. K. I. Gross, "The Plancherel transform on the nilpotent part of G_2 and some application to the representation theory of G_2," *Trans. Amer. Math. Soc.*, vol. 132, pp. 411–446, 1968.
112. V. W. Guillemin and S. Sternberg, "Remarks on a paper of Hermann," *Trans. Amer. Math. Soc.*, vol. 130, pp. 110–116, 1968.
113. W. Hahn, *Stability of Motion*. Berlin: Springer-Verlag, 1967.
114. O. Hájek, "Bilinear control: rank-one inputs," *Funkcial. Ekvac.*, vol. 34, no. 2, pp. 355–374, 1991.
115. B. C. Hall, *An Elementary Introduction to Groups and Representations*. New York: Springer-Verlag, 2003.
116. M. Hall, "A basis for free Lie rings," *Proc. Amer. Math. Soc.*, vol. 1, pp. 575–581, 1950.
117. P. Hartman, *Ordinary Differential Equations*. New York: Wiley, 1964; 1973 [paper, author's reprint].
118. G. W. Haynes, "Controllability of nonlinear systems," Natl. Aero. & Space Admin., Tech. Rep. NASA CR-456, 1965.
119. G. W. Haynes and H. Hermes, "Nonlinear controllability via Lie theory," *SIAM J. Control Opt.*, vol. 8, pp. 450–460, 1970.
120. S. Helgason, *Differential Geometry and Symmetric Spaces*. New York: Academic Press,, 1972.
121. P. Henrici, *Applied and Computational Complex Analysis*. New York: Wiley, 1988, vol. I: Power Series-Integration-Conformal Mapping-Location of Zeros.
122. R. Hermann, "On the accessibility problem in control theory," in *Int. Symp. on Nonlinear Differential Equations and Nonlinear Mechanics*, J. LaSalle and S. Lefschetz, Eds. New York: Academic Press, 1963.
123. R. Hermann, *Differential Geometry and the Calculus of Variations*. New York: Academic Press, 1968, Second edition: Math Sci Press, 1977.
124. R. Hermann, "The formal linearization of a semisimple Lie algebra of vector fields about a singular point," *Trans. Amer. Math. Soc.*, vol. 130, pp. 105–109, 1968.
125. H. Hermes, "Lie algebras of vector fields and local approximation of attainable sets," *SIAM J. Control Optim.*, vol. 16, no. 5, pp. 715–727, 1978.
126. R. Hide, A. C. Skeldon, and D. J. Acheson, "A study of two novel self-exciting single-disk homopolar dynamos: Theory," *Proc. Royal Soc. (London)*, vol. A 452, pp. 1369–1395, 1996.
127. J. Hilgert, K. H. Hofmann, and J. D. Lawson, *Lie Groups, Convex Cones, and Semigroups*. Oxford: Clarendon Press, 1989.
128. J. Hilgert and K.-H. Neeb, *Lie Semigroups and their Applications*. Berlin: Springer-Verlag: Lec. Notes in Math. **1552**, 1993.
129. G. Hochschild, *The Structure of Lie Groups*. San Francisco: Holden-Day, 1965.
130. L. Hörmander, "Hypoelliptic second order differential equations," *Acta Mathematica*, vol. 119, pp. 147–171, 1967.
131. R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1985, 1993.
132. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1991.

133. G. M. Huang, T. J. Tarn, and J. W. Clark, "On the controllability of quantum-mechanical systems," *J. Math. Phys.*, vol. 24, no. 11, pp. 2608–2618, 1983.
134. T. Huillet, A. Monin, and G. Salut, "Minimal realizations of the matrix transition lie group for bilinear control systems: explicit results," *Syst. Control Lett.*, vol. 9, no. 3, pp. 267–274, 1987.
135. J. E. Humphreys, *Linear Algebraic Groups*. New York: Springer-Verlag, 1975.
136. L. R. Hunt, *Controllability of Nonlinear Hypersurface Systems*. Amer. Math. Soc., 1980, pp. 209–224.
137. L. R. Hunt, "n-dimensional controllability with $(n-1)$ controls," *IEEE Trans. Automat. Control*, vol. AC-27, no. 1, pp. 113–117, 1982.
138. N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*. Amsterdam North-Holland/Kodansha, 1981.
139. A. Isidori, "Direct construction of minimal bilinear realizations from nonlinear input–output maps." *IEEE Trans. Automat. Control*, vol. AC-18, pp. 626–631, 1973.
140. A. Isidori, *Nonlinear Control Systems*, 3rd ed. London: Springer-Verlag, 1995.
141. A. Isidori and A. Ruberti, "Realization theory of bilinear systems." in *Geometric Methods in System Theory*. Dordrecht, The Netherlands: D. Reidel, 1973, pp. 83–110.
142. A. Isidori and A. Ruberti, "Time-varying bilinear systems," in *Variable Structure Systems with Application to Economics and Biology*. Berlin: Springer-Verlag: Lecture Notes in Econom. and Math. Systems **111**, 1975, pp. 44–53.
143. N. Jacobson, *Lie Algebras*, ser. Tracts in Math., No. 10. New York: Wiley, 1962.
144. M. Johansson, *Piecewise Linear Control Systems*. Berlin: Springer-Verlag: Lec. Notes in Control and Information Sciences **284**, 2003.
145. I. Joó and N. M. Tuan, "On controllability of bilinear systems. II. Controllability in two dimensions," *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, vol. 35, pp. 217–265, 1992.
146. K. Josić and R. Rosenbaum, "Unstable solutions of nonautonomous linear differential equations," *SIAM Rev.*, vol. 50, no. 3, pp. 570–584, 2008.
147. V. Jurdjevic, *Geometric Control Theory*. Cambridge, UK: Cambridge University Press, 1996.
148. V. Jurdjevic and I. Kupka, "Control systems subordinated to a group action: Accessibility," *J. Differ. Equ.*, vol. 39, pp. 186–211, 1981.
149. V. Jurdjevic and J. Quinn, "Controllability and stability." *J. Differ. Equ.*, vol. 28, pp. 381–389, 1978.
150. V. Jurdjevic and G. Sallet, "Controllability properties of affine systems," *SIAM J. Control Opt.*, vol. 22, pp. 501–508, 1984.
151. V. Jurdjevic and H. J. Sussmann, "Controllability on Lie groups." *J. Differ. Equ.*, vol. 12, pp. 313–329, 1972.
152. T. Kailath, *Linear Systems*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.
153. G. Kallianpur and C. Striebel, "A stochastic differential equation of Fisk type for estimation and nonlinear filtering problems," *SIAM J. Appl. Math.*, vol. 21, pp. 61–72, 1971.
154. R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in Mathematical System Theory*. New York: McGraw-Hill, 1969.
155. N. Kalouptsidis and J. Tsinias, "Stability improvement of nonlinear systems by feedback," *IEEE Trans. Automat. Control*, vol. AC-28, pp. 364–367, 1984.
156. H. Khalil, *Nonlinear Systems*. Upper Saddle River, N.J.: Prentice Hall, 2002.
157. N. Khaneja, S. J. Glaser, and R. Brockett, "Sub-Riemannian geometry and time optimal control of three spin systems: quantum gates and coherence transfer," *Phys. Rev. A (3)*, vol. 65, no. 3, part A, pp. 032 301, 11, 2002.
158. A. Y. Khapalov and R. R. Mohler, "Reachability sets and controllability of bilinear time-invariant systems: A qualitative approach," *IEEE Trans. Automat. Control*, vol. 41, no. 9, pp. 1342–1346, 1996.

159. T. Kimura, *Introduction to Prehomogeneous Vector Spaces*. Providence, RI: American Mathematical Society, 2002, Trans. of Japanese edition: Tokyo, Iwanami Shoten, 1998.
160. M. K. Kinyon and A. A. Sagle, "Quadratic dynamical systems and algebras," *J. Differ. Equ.*, vol. 117, no. 1, pp. 67–126, 1995.
161. H. Kraft, "Algebraic automorphisms of affine space." in *Topological Methods in Algebraic Transformation Groups: Proc. Conf. Rutgers University*, H. Kraft, T. Petrie, and G. W. Schwartz, Eds. Boston, Mass.: Birkhäuser, 1989.
162. L. Kramer, "Two-transitive Lie groups," *J. Reine Angew. Math.*, vol. 563, pp. 83–113, 2003.
163. N. N. Krasovskii, *Stability of Motion*. Stanford, Calif.: Stanford University Press, 1963.
164. A. J. Krener, "On the equivalence of control systems and the linearization of nonlinear systems," *SIAM J. Control Opt.*, vol. 11, pp. 670–676, 1973.
165. A. J. Krener, "A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems," *SIAM J. Control Opt.*, vol. 12, pp. 43–52, 1974.
166. A. J. Krener, "Bilinear and nonlinear realizations of input–output maps." *SIAM J. Control Optim.*, vol. 13, pp. 827–834, 1975.
167. J. Kučera, "Solution in large of control problem x′ = (A(1-u)+Bu)x." *Czechoslovak Math. J*, vol. 16 (91), pp. 600–623, 1966, see "The control problem x′ = (A(1 − u) + Bu)x: a comment on an article by J. Kučera".
168. J. Kučera, "Solution in large of control problem x′ = (Au +Bv)x," *Czechoslovak Math. J*, vol. 17 (92), pp. 91–96, 1967.
169. J. Kučera, "On accessibility of bilinear systems." *Czechoslovak Math. J*, vol. 20, pp. 160–168, 1970.
170. M. Kuranishi, "On everywhere dense embedding of free groups in Lie groups," *Nagoya Math. J.*, vol. 2, pp. 63–71, 1951.
171. A. G. Kušnirenko, "An analytic action of a semisimple lie group in a neighborhood of a fixed point is equivalent to a linear one (russian)," *Funkcional. Anal. i Priložen [trans: Funct. Anal. Appl.]*, vol. 1, no. 2, pp. 103–104, 1967.
172. G. Lafferriere and H. J. Sussmann, "Motion planning for controllable systems without drift," in *Proc. IEEE Conf. Robotics and Automation, Sacramento, CA, April 1991*. Piscataway, N.J.: IEEE Pubs., 1991, pp. 1148–1153.
173. J. Lawson, "Geometric control and Lie semigroup theory," in *Differential Geometry and Control*, G. Ferreyra, R. Gardner, H. Hermes, and H. Sussmann, Eds. Providence: AMS, 1999, pp. 207–221.
174. J. D. Lawson and D. Mittenhuber, "Controllability of Lie systems," in *Contemporary Trends in Nonlinear Geometric Control Theory and Its Applications (México City, 2000)*. Singapore: World Scientific, 2002, pp. 53–76.
175. U. Ledzewicz and H. Schättler, "Analysis of a class of optimal control problems arising in cancer chemotherapy," in *Proc. Joint Automatic Control Conf., Anchorage, Alaska, May 8-10, 2002*. AACC, 2002, pp. 3460–3465.
176. U. Ledzewicz and H. Schättler, "Optimal bang-bang controls for a two-compartment model in cancer chemotherapy," *J. Optim. Theory Appl.*, vol. 114, no. 3, pp. 609–637, 2002.
177. E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*. New York: John Wiley & Sons Inc., 1967.
178. E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, 2nd ed. Melbourne, FL: Robert E. Krieger Publishing Co. Inc., 1986.
179. K. Y. Lee, "Optimal bilinear control theory applied to pest management," in *Recent Developments in Variable Structure Systems, Economics and Biology. (Proc. U.S. – Italy Seminar, Taormina, Sicily, 1977)*, R. R. Mohler and A. Ruberti, Eds. Berlin: Springer-Verlag. 1978, pp. 173–188.
180. F. S. Leite and P. E. Crouch, "Controllability on classical lie groups," *Math. Control Sig. Sys.*, vol. 1, pp. 31–42, 1988.

181. N. L. Lepe, "Geometric method of investigation of the controllability of second-order bilinear systems," *Avtomat. i Telemekh.*, no. 11, pp. 19–25, 1984.
182. D. Liberzon, *Switching in Systems and Control*. Boston, MA: Birkhäuser, 1993.
183. D. Liberzon, "Lie algebras and stability of switched nonlinear systems," in *Unsolved Problems in Mathematical Systems & Control Theory*, V. D. Blondel and A. Megretski, Eds. Princeton, NJ: Princeton University Press, 2004, ch. 6.4, pp. 203–207.
184. D. Liberzon, J. P. Hespanha, and A. S. Morse, "Stability of switched systems: a Lie-algebraic condition," *Syst. Control Lett.*, vol. 37, no. 3, pp. 117–122, 1999.
185. W. Lin and C. I. Byrnes, "KYP lemma, state feedback and dynamic output feedback in discrete-time bilinear systems," *Syst. Control Lett.*, vol. 23, no. 2, pp. 127–136, 1994.
186. W. Liu, "An approximation algorithm for nonholonomic systems," *SIAM J. Control Opt.*, vol. 35, pp. 1328–1365, 1997.
187. E. S. Livingston, "Linearization of analytic vector fields with a common critical point," Ph.D. Dissertation, Washington University, St. Louis, Missouri, 1982.
188. E. S. Livingston and D. L. Elliott, "Linearization of families of vector fields." *J. Differ. Equ.*, vol. 55, pp. 289–299, 1984.
189. C. Lobry, "Contrôllabilité des systèmes non linéaires," *SIAM J. Control Opt.*, vol. 8, pp. 573–605, 1970.
190. C. Lobry, "Controllability of nonlinear systems on compact manifolds," *SIAM J. Control*, vol. 12, pp. 1–4, 1974.
191. H. Logemann and E. P. Ryan, "Asymptotic behaviour of nonlinear systems," *Amer. Math. Monthly*, vol. 11, pp. 864–889, December 2004.
192. R. Longchamp, "Stable feedback control of bilinear systems," *IEEE Trans. Automat. Control*, vol. AC-25, pp. 302–306, April 1980.
193. E. N. Lorenz, "Deterministic nonperiodic flow," *J. Atmos. Sciences*, vol. 20, pp. 130–148, 1963.
194. R. Luesink and H. Nijmeijer, "On the stabilization of bilinear systems via constant feedback," *Linear Algebra Appl.*, vol. 122/123/124, pp. 457–474, 1989.
195. W. Magnus, "On the exponential solution of differential equations for a linear operator," *Comm. Pure & Appl. Math.*, vol. 7, pp. 649–673, 1954.
196. A. Malcev, "On the theory of lie groups in the large," *Rec. Math. [Mat. Sbornik]*, vol. 16 (58), pp. 163–190, 1945.
197. G. Marchesini and S. K. Mitter, Eds., *Mathematical Systems Theory*. Berlin: Springer-Verlag: Lec. Notes in Econ. and Math. Systems **131**, 1976, proceedings of International Symposium, Udine, Italy, 1975.
198. M. Margaliot, "Stability analysis of switched systems using variational principles: an introduction'," *Automatica—J. IFAC*, vol. 42, no. 12, pp. 2059–2077, 2006. [Online]. Available: http://www.eng.tau.ac.il/~michaelm/survey.pdf
199. M. Margaliot, "On the reachable set of nonlinear control systems with a nilpotent Lie algebra," in *Proc. 9th European Control Conf. (ECC'07)*, Kos, Greece, 2007, pp. 4261–4267. [Online]. Available: www.eng.tau.ac.il/~michaelm
200. L. Markus, "Quadratic differential equations and non-associative algebras," in *Contributions to the Theory of Nonlinear Oscillations, Vol. V*. Princeton, N.J.: Princeton Univ. Press, 1960, pp. 185–213.
201. L. Markus, "Exponentials in algebraic matrix groups," *Advances in Math.*, vol. 11, pp. 351–367, 1973.
202. P. Mason, U. Boscain, and Y. Chitour, "Common polynomial Lyapunov functions for linear switched systems," *SIAM J. Control Opt.*, vol. 45, pp. 226–245, 2006. [Online]. Available: http://epubs.siam.org/fulltext/SICON/volume-45/61314.pdf
203. D. Q. Mayne and R. W. Brockett, Eds., *Geometric Methods in System Theory*. Dordrecht, The Netherlands: D. Reidel, 1973.

204. R. Mohler and A. Ruberti, Eds., *Theory and Applications of Variable Structure Systems*. New York: Academic Press, 1972.
205. R. R. Mohler, *Controllability and Optimal Control of Bilinear Systems*. Englewood Cliffs, N.J.: Prentice-Hall, 1970.
206. R. R. Mohler, *Bilinear Control Processes*. New York: Academic Press, 1973.
207. R. R. Mohler, *Nonlinear Systems: Volume II, Applications to Bilinear Control*. Englewood Cliffs, N.J.: Prentice-Hall, 1991.
208. R. R. Mohler and R. E. Rink, "Multivariable bilinear system control." in *Advances in Control Systems*, C. T. Leondes, Ed. New York: Academic Press, 1966, vol. 2.
209. R. R. Mohler and A. Ruberti, Eds., *Recent Developments in Variable Structure Systems, Economics and Biology. (Proc. U.S. – Italy Seminar, Taormina, Sicily, 1977)*. Berlin: Springer-Verlag, 1978.
210. R. Mohler and W. Kolodziej, "An overview of bilinear system theory and applications." *IEEE Trans. Syst., Man Cybern.*, vol. SMC-10, no. 10, pp. 683–688, 1980.
211. C. Moler and C. F. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later," *SIAM Rev.*, vol. 45, pp. 3–49, 2003.
212. D. Montgomery and H. Samelson, "Transformation groups of spheres," *Ann. of Math.*, vol. 44, pp. 454–470, 1943.
213. D. Montgomery and L. Zippin, *Topological Transformation Groups*. New York: Interscience Publishers, 1955, reprinted by Robert E. Krieger, 1974.
214. C. B. Morrey, "The analytic embedding of abstract real-analytic manifolds," *Ann. of Math. (2)*, vol. 68, pp. 159–201, 1958.
215. S. Nonoyama, "Algebraic aspects of bilinear systems and polynomial systems," D.Sc. Dissertation, Washington University, St. Louis, Missouri, August 1976.
216. O. A. Oleinik and E. V. Radkevich, "On local smoothness of generalized solutions and hypoellipticity of second order differential equations," *Russ. Math. Surveys*, vol. 26, pp. 139–156, 1971.
217. A. I. Paolo d'Allessandro and A. Ruberti, "Realization and structure theory of bilinear dynamical systems," *SIAM J. Control Opt.*, vol. 12, pp. 517–535, 1974.
218. M. A. Pinsky, *Lectures on Random Evolution*. Singapore: World Scientific, 1991.
219. H. Poincaré, *Œuvres*. Paris: Gauthier-Villars, 1928, vol. I.
220. L. Pontryagin, V. Boltyanskii, R. Gamkrelidze, and E. Mishchenko, *The Mathematical Theory of Optimal Processes*. New York: MacMillan, 1964.
221. C. Procesi, "The invariant theory of $n \times n$ matrices," *Adv. in Math.*, vol. 19, no. 3, pp. 306–381, 1976.
222. S. M. Rajguru, M. A. Ifediba, and R. D. Rabbitt, "Three-dimensional biomechanical model of benign paroxysmal positional vertigo," *Ann. Biomed. Eng.*, vol. 32, no. 6, pp. 831–846, 2004. [Online]. Available: http://dx.doi.org/10.1023/B:ABME.0000030259.41143.30
223. C. Reutenauer, *Free Lie Algebras*. London: Clarendon Press, 1993.
224. R. E. Rink and R. R. Mohler, "Completely controllable bilinear systems." *SIAM J. Control Optim.*, vol. 6, pp. 477–486, 1968.
225. W. Rossmann, *Lie Groups: An Introduction Through Linear Groups*. Oxford: Oxford University Press, 2003. Corrected printing: 2005.
226. H. Rubenthaler, "Formes réeles des espaces préhomogènes irrèducibles de type parabolique," *Ann. Inst. Fourier (Grenoble)*, pp. 1–38, 1986.
227. A. Ruberti and R. R. Mohler, Eds., *Variable Structure Systems with Applications to Econonomics and Biology. (Proceedings 2nd U.S. – Italy Seminar, Portland, Oregon, 1974)*. Berlin: Springer-Verlag: Lec. Notes in Econ. and Math. Systems **111**, 1975.
228. W. J. Rugh, *Nonlinear System Theory: The Volterra/Wiener Approach*. Baltimore: Johns Hopkins University Press, 1981. [Online]. Available: http://www.ece.jhu.edu/~rugh/volterra/book.pdf
229. Y. L. Sachkov, "On invariant orthants of bilinear systems," *J. Dynam. Control Syst.*, vol. 4, no. 1, pp. 137–147, 1998.

230. Y. L. Sachkov, "Controllability of affine right-invariant systems on solvable Lie groups," *Discrete Math. Theor. Comput. Sci.*, vol. 1, pp. 239–246, 1997.
231. Y. L. Sachkov, "On positive orthant controllability of bilinear systems in small codimensions." *SIAM J. Control Optim.*, vol. 35, no. 1, pp. 29–35, January 1997.
232. Y. L. Sachkov, "Controllability of invariant systems on Lie groups and homogeneous spaces," *J. Math. Sci. (New York)*, vol. 100, pp. 2355–2477, 2000.
233. L. A. B. San Martin, "On global controllability of discrete-time control systems," *Math. Control Signals Systems*, vol. 8, no. 3, pp. 279–297, 1995.
234. M. Sato and T. Shintani, "On zeta functions associated with prehomogeneous vector spaces," *Proc. Nat. Acad. Sci. USA*, vol. 69, no. 2, pp. 1081–1082, May 1972.
235. L. I. Schiff, *Quantum Mechanics*. New York: McGraw-Hill, 1968.
236. S. G. Schirmer and A. I. Solomon, "Group-theoretical aspects of control of quantum systems," in *Proceedings 2nd Int'l Symposium on Quantum Theory and Symmetries (Cracow, Poland, July 2001)*. Singapore: World Scientific, 2001. [Online]. Available: http://cam.qubit.org/users/sonia/research/papers/2001Cracow.pdf
237. J. L. Sedwick, "The equivalence of nonlinear and bilinear control systems," Sc.D. Dissertation, Washington University, St. Louis, Missouri, 1974.
238. J. L. Sedwick and D. L. Elliott, "Linearization of analytic vector fields in the transitive case," *J. Differ. Equ.*, vol. 25, pp. 377–390, 1977.
239. P. Sen, "On the choice of input for observability in bilinear systems." *IEEE Trans. Automat. Control*, vol. AC-26, no. 2, pp. 451–454., 1981.
240. J.-P. Serre, *Lie Algebras and Lie Groups*. New York: W.A. Benjamin, Inc., 1965.
241. R. Shorten, F. Wirth, O. Mason, K. Wulf, and C. King, "Stability criteria for switched and hybrid systems," *SIAM Rev.*, vol. 49, no. 4, pp. 545–592, December 2007.
242. H. Sira-Ramirez, "Sliding motions in bilinear switched networks," *IEEE Trans. Circuits Syst.*, vol. CAS-34, no. 8, pp. 919–933, 1987.
243. M. Slemrod, "Stabilization of bilinear control systems with applications to nonconservative problems in elasticity." *SIAM J. Control Optim.*, vol. 16, no. 1, pp. 131–141, January 1978.
244. E. D. Sontag, *Polynomial Response Maps*. Berlin: Springer-Verlag: Lecture Notes in Control and Information Science **13**, 1979.
245. E. D. Sontag, "Realization theory of discrete-time nonlinear systems: Part I — the bounded case," *IEEE Trans. Circuits Syst.*, vol. CAS-26, no. 4, pp. 342–356, 1979.
246. E. D. Sontag, "On the observability of polynomial systems, I: Finite time problems." *SIAM J. Control Optim.*, vol. 17, pp. 139–151, 1979.
247. E. D. Sontag, "A remark on bilinear systems and moduli spaces of instantons," *Systems Control Lett.*, vol. 9, pp. 361–367, 1987.
248. E. D. Sontag, "A Chow property for sampled bilinear systems," in *Analysis and Control of Nonlinear Systems*, C. Byrnes, C. Martin, and R. Saeks, Eds. Amsterdam: North Holland, 1988, pp. 205–211.
249. E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, 2nd ed. New York: Springer-Verlag, 1998. [Online]. Available: http://www.math.rutgers.edu/~sontag/FTP_DIR/sontag_mathematical_control_theory_springer98.pdf
250. E. D. Sontag, Y. Wang, and A. Megretski, "Input classes for identification of bilinear systems," *IEEE Trans. Automat. Control*, to appear.
251. P. Stefan and J. B. Taylor, "A remark on a paper of D. L. Elliott," *J. Differ. Equ.*, vol. 15, pp. 210–211, 1974.
252. S. Sternberg, *Lectures on Differential Geometry*. Englewood Cliffs, N.J.: Prentice-Hall, 1964.
253. R. L. Stratonovich, "A new representation for stochastic integrals and equations," *SIAM J. Control*, vol. 4, pp. 362–371, 1966.
254. H. K. Struemper, "Motion control for nonholonomic systems on matrix lie groups," Sc.D. Dissertation, University of Maryland, College Park, Maryland, 1997.

255. B. Sturmfels, "What is a Gröbner basis?" *Notices Amer. Math. Soc.*, vol. 50, no. 2, pp. 2–3, November 2005.
256. H. J. Sussmann, "The "bang-bang" problem for certain control systems in GL(n,R)." *SIAM J. Control Optim.*, vol. 10, no. 3, pp. 470–476, August 1972.
257. H. J. Sussmann, "The control problem $x' = (A(1-u) + Bu)x$: a comment on an article by J. Kucera," *Czech. Math. J.*, vol. 22, pp. 423–426, 1972.
258. H. J. Sussmann, "Orbits of families of vector fields and integrability of distributions," *Trans. Amer. Math. Soc.*, vol. 180, pp. 171–188, 1973.
259. H. J. Sussmann, "An extension of a theorem of Nagano on transitive Lie algebras," *Proc. Amer. Math. Soc.*, vol. 45, pp. 349–356, 1974.
260. H. J. Sussmann, "On the number of directions necessary to achieve controllability." *SIAM J. Control Optim.*, vol. 13, pp. 414–419, 1975.
261. H. J. Sussmann, "Semigroup representations, bilinear approximation of input-output maps, and generalized inputs." in *Mathematical Systems Theory*, G. Marchesini and S. K. Mitter, Eds. Berlin: Springer-Verlag, 1975, pp. 172–191.
262. H. J. Sussmann, "Minimal realizations and canonical forms for bilinear systems." *J. Franklin Inst.*, vol. 301, pp. 593–604, 1976.
263. H. J. Sussmann, "On generalized inputs and white noise," in *Proc. 1976 IEEE Conf. Decis. Contr.*, vol. 1. New York: IEEE Pubs., 1976, pp. 809–814.
264. H. J. Sussmann, "On the gap between deterministic and stochastic differential equations," *Ann. Probab.*, vol. 6, no. 1, pp. 19–41, 1978.
265. H. J. Sussmann, "Single-input observability of continuous-time systems." *Math. Systems Theory*, vol. 12, pp. 371–393, 1979.
266. H. J. Sussmann, "A product expansion for the Chen series," in *Theory and Applications of Nonlinear Control Systems*, C. I. Byrnes and A. Lindquist, Eds. North Holland: Elsevier Science Publishers B. V., 1986, pp. 323–335.
267. H. J. Sussmann, "A continuation method for nonholonomic path-finding problems," in *Proc. 32nd IEEE Conf. Decis. Contr.* New York: IEEE Pubs., 1993, pp. 2718–2723. [Online]. Available: https://www.math.rutgers.edu/~sussmann/papers/cdc93-continuation.ps.gz
268. H. J. Sussmann, "Geometry and optimal control," in *Mathematical Control Theory*. New York: Springer-Verlag, 1999, pp. 140–198. [Online]. Available: http://math.rutgers.edu/~sussmann/papers/brockettpaper.ps.gz[includescorrections]
269. H. J. Sussmann and V. Jurdjevic, "Controllability of nonlinear systems." *J. Differ. Equ.*, vol. 12, pp. 95–116, 1972.
270. H. J. Sussmann and W. Liu, "Limits of highly oscillatory controls and the approximation of general paths by admissible trajectories," in *Proc. 30th IEEE Conf. Decis. Contr.* New York: IEEE Pubs., December 1991, pp. 437–442.
271. H. J. Sussmann and W. Liu, "Lie bracket extensions and averaging: the single-bracket case," in *Nonholonomic Motion Planning*, ser. ISECS, Z. X. Li and J. F. Canny, Eds. Boston: Kluwer Acad. Publ., 1993, pp. 109–148.
272. H. J. Sussmann and J. Willems, "300 years of optimal control: from the brachystochrone to the maximum principle," *IEEE Contr. Syst. Magazine*, vol. 17, no. 3, pp. 32–44, June 1997.
273. M. Suzuki, "On the convergence of exponential operators—the Zassenhaus formula, BCH formula and systematic approximants," *Comm. Math. Phys.*, vol. 57, no. 3, pp. 193–200, 1977.
274. T.-J. Tarn, D. L. Elliott, and T. Goka, "Controllability of discrete bilinear systems with bounded control." *IEEE Trans. Automat. Control*, vol. 18, pp. 298–301, 1973.
275. T.-J. Tarn and S. Nonoyama, "Realization of discrete-time internally bilinear systems." in *Proc. 1976 IEEE Conf. Decis. Contr.* New York: IEEE Pubs., 1976, pp. 125–133;.
276. T.-J. Tarn and S. Nonoyama, "An algebraic structure of discrete-time biaffine systems," *IEEE Trans. Automat. Control*, vol. AC-24, no. 2, pp. 211–221, April 1979.

277. J. Tits, *Tabellen zu den Einfachen Lie gruppen und ihren Darstellungen*. Springer-Verlag: Lec. Notes Math. **40**, 1967.
278. M. Torres-Torriti and H. Michalska, "A software package for Lie algebraic computations," *SIAM Rev.*, vol. 47, no. 4, pp. 722–745, 2005.
279. G. Turinici and H. Rabitz, "Wavefunction controllability for finite-dimensional bilinear quantum systems," *J. Phys. A*, vol. 36, no. 10, pp. 2565–2576, 2003.
280. V. I. Utkin, "Variable structure systems with sliding modes," *IEEE Trans. Automat. Control*, vol. AC-22, no. 2, pp. 212–222, 1977.
281. V. J. Utkin, *Sliding Modes in Control and Optimization*. Berlin: Springer-Verlag, 1992, translated and revised from the 1981 Russian original.
282. V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations.*, 2nd ed., ser. GTM, Vol. 102. Berlin and New York: Springer-Verlag, 1988.
283. W. C. Waterhouse, "Re: n-dimensional group actions on R^n," personal communication, May 7 2003.
284. J. Wei and E. Norman, "Lie algebraic solution of linear differential equations." *J. Math. Phys.*, vol. 4, pp. 575–581, 1963.
285. J. Wei and E. Norman, "On global representations of the solutions of linear differential equations as products of exponentials." *Proc. Amer. Math. Soc.*, vol. 15, pp. 327–334, 1964.
286. R. M. Wilcox, "Exponential operators and parameter differentiation in quantum physics." *J. Math. Phys.*, vol. 8, pp. 962–982, 1967.
287. D. Williamson, "Observation of bilinear systems with application to biological control." *Automatica—J. IFAC*, vol. 13, pp. 243–254, 1977.
288. R. L. Winslow, *Theoretical Foundations of Neural Modeling: BME 580.681*. Baltimore: Unpublished, 1992. [Online]. Available: http://www.ccbm.jhu.edu/doc/courses/BME_580.681/Book/book.ps
289. F. Wirth, "A converse Lyapunov theorem for linear parameter-varying and linear switching systems," *SIAM J. Control Optim.*, vol. 44, no. 1, pp. 210–239, 2005.

This page intentionally left blank

Index

D.L. Elliott, *Bilinear Control Systems,* Applied Mathematical Sciences 169,
DOI 10.1007/978-1-4020-9613-6_14,