Chapter 12
Control of Ensembles

The purpose of this chapter is to provide an introduction to the emerging field of ensemble control for linear systems, i.e., the control of families of linear systems. Ensemble control refers to the task of controlling a large, potentially infinite, number of states, or systems, using a single-input function or a single-feedback controller. Thus, controlling ensembles is very much at the core of a robust theory of networks of systems. The study of ensembles is motivated from quite different applications.

- A classical example from physics concerns the conversion of heat into work by control of the heat flow and volume. Here the node systems are the gas molecules whose dynamics are described via the formalism of statistical mechanics. Of course, the sheer magnitude of the system, consisting roughly of $N \simeq 6 \times 10^{23}$ molecules per mole, calls for a thermodynamical or statistical mechanics approach rather than ab initio calculations.
- Quantum control of weakly coupled identical particles in nuclear magnetic resonance spectroscopy is another interesting area that has attracted much interest recently. Here the goal is to control a possibly large system of spins by applying short pulses of an electromagnetic field. The system is described by a controlled Liouville–von Neumann master equation evolving on density operators or Lie groups of unitary matrices. A closely related question concerns the control of parameterized families of systems using open-loop controls that are independent of the parameters. Open-loop control has recently acquired popularity in quantum control through the work of Li and Khaneja (2006), for example.
- The task of controlling probability distributions in the state space for a finite-dimensional control system leads to the control of Liouville transport equations or, more generally, of Fokker–Planck equations.

P.A. Fuhrmann, U. Helmke, *The Mathematics of Networks of Linear Systems*, Universitext, DOI 10.1007/978-3-319-16646-9_12

- Another area of interest is the control of parameter-varying systems using parameter-independent open-loop or closed-loop controllers. As a specific instance we mention the so-called **blending problem** from the late 1970s, which seeks to find parameter-independent feedback control laws that stabilize a family of linear systems.

In all these areas the question arises of how to approximately control a family of systems or of state variables. Thus the key ideas behind studying ensembles of linear systems are the same as those for studying large-scale systems. For complex dynamical systems the goals of controlling the entire collection of states is often asking for too much and needs to be replaced by more realistic goals, such as controlling, for example, the mean value or covariance of the state variables. This happens, for example, in the control of open quantum systems where the state variables are defined by density operators that describe an average of state variables. It also happens in daily life experience, for example, in cooking a meal in an oven. Here the interesting task is not so much to control the temperature distribution in the oven but rather to create an average temperature profile that yields the desired result. The same issue arises in motion control problems for infinite platoons of vehicles, as in the work by Rogge and Aeyels (2008), or for spatiotemporal systems described by partial differential equations (PDEs), on which see Bamieh, Paganini and Dahleh (2002). Using Fourier-transform techniques, spatially invariant control systems can be identified with parameter-dependent families of linear systems. This opens the door to applications of a variety of different approaches to distributed large-scale systems. Parameter-dependent systems can also be regarded as infinite-dimensional systems defined on suitable Banach or Hilbert spaces of functions. This brings about many opportunities for interactions between functional analysis and parametric systems. We also mention the theory of systems over rings as a systematic algebraic approach to analyzing parameter-dependent systems.

In this chapter we focus on the specific task of finding **open-loop controls** that steer a system from a family of initial states to another family of terminal states using a single open-loop control function that is independent of the parameters. We then illustrate how this question arises in a number of different applications, i.e., for controlling PDEs, controlling spatially-invariant systems such as platoons or to the robust control of networks of linear systems. We turn now to a more detailed description of such issues and begin to explain their mutual relationships.

12.1 Control of Parametric Families of Systems

The starting point for our analysis of ensembles of linear systems are parameter-dependent linear systems of the form

$$\frac{\partial}{\partial t}x(t,\theta) = A(\theta)x(t,\theta) + B(\theta)u(t), \quad x(0,\theta) = 0. \tag{12.1}$$

Of course, one could equally well consider families of discrete-time systems

$$x(t+1,\theta) = A(\theta)x(t,\theta) + B(\theta)u(t), \quad x(0,\theta) = 0,$$

and our results apply to this case, too. For simplicity we assume that the system matrices $A(\theta) \in \mathbb{R}^{n\times n}$ and $B(\theta) \in \mathbb{R}^{n\times m}$ vary continuously in a compact domain $\mathbf{P}$ of parameters θ in Euclidean space $\mathbb{R}^d$. The analysis of such families of linear systems can be carried out in several directions. A straightforward issue to begin with is the search for parameter-dependent controls that steer the systems from a family of initial states to a family of desired terminal states. A restriction here might be on the degree of continuity or smoothness in the parameters that is imposed on the controls. For instance, if the system matrices depend polynomially on a parameter, it may be desirable that the same be true for the feedback controllers and input functions. This leads to the control problems for systems over rings that have been discussed intensively in algebraic systems theory during the course of the past four decades. We refer the reader to the early work of, for example, Hazewinkel (1981), Sontag (1976), Tannenbaum (1981), and Conte and Perdon (2000) for further details. Another extreme case of studying families of systems – and this is the scenario that we will study in the remainder of this chapter – is to search for input functions or feedback laws that are independent of the parameters of systems and steer prescribed families of initial and terminal states arbitrarily close to each other. We refer to this as the **ensemble control** problem. Thus ensemble control refers to a specific class of robust control problems, and a priori it is not obvious whether or not such problems can be solved.

Let p and q be integers, with $1 \le p,q \le \infty$. The input to a state operator of (12.1) at time T is $\mathscr{R}_T : L^p([0,T],\mathbb{R}^m) \longrightarrow L^q(\mathbf{P},\mathbb{R}^n)$

$$\mathscr{R}_T(u)(\theta) = \int_0^T e^{(T-s)A(\theta)}B(\theta)u(s)ds.$$

Note that $\mathscr{R}_T$ is an integral operator with continuous kernel $K : \mathbf{P} \times [0,T] \longrightarrow \mathbb{R}^n$,

$$K(\theta,s) = e^{(T-s)A(\theta)}B(\theta).$$

It is a well-known consequence of the assumed continuity of $A(\cdot),B(\cdot)$ on the compact parameter space $\mathbf{P}$ that $\mathscr{R}_T$ defines a bounded linear operator. In fact, $\mathscr{R}_T$ is a compact operator for $1 < p \le \infty$ and $1 \le q < \infty$. Moreover, by the continuity of $A(\cdot),B(\cdot)$, the operator $\mathscr{R}_T$ is compact, even as an operator from $L^p([0,T])$ to $C(\mathbf{P},\mathbb{R}^n)$ and $1 \le p \le \infty$. From the compactness of $\mathscr{R}_T$ we deduce that $\mathscr{R}_T$ has a closed image if and only if $\mathscr{R}_T$ has finite rank. Clearly, $\mathscr{R}_T$ has finite rank for each linear operator of the form $K(\theta,t) = \sum_{j=1}^k \phi_j(t)\psi_j(\theta)$. In particular, this is true if A is parameter independent; however, for general parameter-dependent systems the finite-rank property cannot be expected.

After these functional analytic generalities, we proceed to introduce the notion of ensemble reachability that we are interested in.

Definition 12.1. Let $1 \leq p \leq \infty$. System (12.1) is **uniformly ensemble reachable** if for every continuous function $x^* : \mathbf{P} \longrightarrow \mathbb{R}^n$ and every $\varepsilon > 0$ there exists a control $u \in L^p([0,T], \mathbb{R}^m)$ such that the induced state trajectory satisfies

$$\sup_{\theta \in \mathbf{P}} \|x(T,\theta) - x^*(\theta)\| < \varepsilon. \tag{12.2}$$

Instead of trying to construct controls that achieve the uniform ensemble reachability condition (12.2), one can also search for controls $u(t)$ that minimize the L^q-norms for $1 \leq q \leq \infty$,

$$\left(\int_{\mathbf{P}} \|x(T,\theta) - x^*(\theta)\|^q d\theta \right)^{\frac{1}{q}} < \varepsilon. \tag{12.3}$$

We then say that the system is L^q**-ensemble reachable**. If the conditions in (12.2) or (12.3) are satisfied for $\varepsilon = 0$, then the system is called **exactly ensemble reachable**. Of course, the important point here is the ability to choose the input function independently of the parameter θ, and it is not at all obvious that systems of this kind do in fact exist.

This notion of ensemble reachability is intimately related to standard concepts from infinite-dimensional systems theory, such as approximate reachability and reachability. Let $\mathscr{A} : X \longrightarrow X$ and $\mathscr{B} : U \longrightarrow X$ be bounded linear operators on Banach spaces X and U, respectively. A linear system

$$\dot{x}(t) = \mathscr{A}x(t) + \mathscr{B}u(t) \tag{12.4}$$

is called **approximately reachable** if the reachable set of 0 is dense in X. See Fuhrmann (1972) and Curtain and Zwart (1995) for the (now classical) characterizations of approximate reachability in a Hilbert space via the condition that the reachability operator has a dense image. A result of Trigianni (1975) implies that parameter-dependent linear systems (12.1) are never exactly ensemble reachable. Thus the approximate notions of (uniform or L^q-) ensemble reachability are the only meaningful ones.

The mathematical connection between ensemble reachability and approximate reachability is easy to describe. Explicitly, for uniform ensemble control, let X denote the Banach space of $\mathbb{R}^n$-valued continuous functions on the compact parameter space $\mathbf{P}$, endowed with a supremum norm. Similarly, for L^q-ensemble reachability, choose $X = L^q(\mathbf{P}, \mathbb{R}^n)$. In either case, a continuous family of linear systems $(A(\theta), B(\theta))$ defines a linear system of the form (12.4) on a Banach space X with a **finite-dimensional** space of control values $U = \mathbb{R}^m$. Here

$$\mathscr{A} : X \longrightarrow X, \quad (\mathscr{A}x)(\theta) := A(\theta)x(\theta) \tag{12.5}$$

denotes the bounded linear multiplication operator, while the input operator

$$\mathscr{B} : \mathbb{R}^m \longrightarrow X, \quad (\mathscr{B}u)(\theta) := B(\theta)u \tag{12.6}$$

is defined via an m-tuple of Banach-space elements, i.e., by the columns of $B(\cdot)$.

Proposition 12.2. *Let bounded linear operators $\mathscr{A}$ and $\mathscr{B}$ be defined as in (12.5) and (12.6), respectively. The parameter-dependent system (12.1) is uniformly (or L^q-) ensemble reachable if and only if system (12.4) on the Banach space X is approximately reachable.*

Proof. We focus on uniform ensemble reachability and the Banach space $X = C(\mathbf{P}, \mathbb{R}^n)$; the proof for L^q-ensemble reachability goes mutatis mutandis. Let $t \mapsto x(t) \in X$ denote the unique solution to (12.4), with $x(0) = 0$. Then $t \mapsto x(t;\theta)$ is the unique solution to (12.1), with $x(0;\theta) = 0$ for $\theta \in \mathbb{P}$. The approximate reachability of (12.4) then says that for the continuous function $x^* : \mathbf{P} \longrightarrow \mathbb{R}^n$ and $\varepsilon > 0$ there exists $T > 0$ such that $\|x(T) - x^*\| = \sup_{\theta \in \mathbf{P}} \|x(T,\theta) - x^*(\theta)\|$. But this is simply the condition for uniform ensemble reachability. ■

Thus the parameter-dependent system (12.1) is uniformly ensemble reachable if and only if the infinite-dimensional system (12.4) is approximately reachable. In the same way, by replacing the Banach space X with the Hilbert space $H = L^2(\mathbf{P}, \mathbb{R}^n)$, one concludes that the L^2-ensemble reachability of (12.1) becomes equivalent to the approximate reachability of the infinite-dimensional system (12.4). Unfortunately, the conditions for approximate reachability stated in Curtain and Zwart (1995); Jacob and Partington (2006) depend on an explicit knowledge of a Riesz basis of eigenvectors for the Hilbert-space operator $\mathscr{A}$. However, except for trivial cases where, for example, $A(\theta)$ has constant eigenvalues, the multiplication operator $\mathscr{A}$ defined by $A(\theta)$ does not have a point spectrum, and therefore the spectral conditions in Curtain and Zwart (1995) are not satisfied here. In the next section, we will explain how such difficulties can be avoided using tools from complex approximation theory.

12.2 Uniform Ensemble Reachability

Next we provide necessary, as well as sufficient, conditions for the uniform ensemble reachability of linear systems (12.1). These conditions are true, verbatim, for discrete-time systems as well. Let

$$(zI - A(\theta))^{-1}B(\theta) = N_\theta(z)D_\theta(z)^{-1}$$

be a right coprime factorization by a rectangular polynomial matrix $N_\theta(z) \in \mathbb{R}^{n\times m}[z]$ and a nonsingular polynomial matrix $D_\theta(z) \in \mathbb{R}^{m\times m}[z]$. We first state the necessary conditions for uniform ensemble reachability.

Proposition 12.3 (Necessary Conditions). *Let $\boldsymbol{P}$ be a subset of $\mathbb{R}^d$ such that the interior points of $\boldsymbol{P}$ are dense in $\boldsymbol{P}$. Assume that the family of linear systems $(A(\theta),B(\theta))_{\theta\in\boldsymbol{P}}$ is uniformly ensemble reachable. Then the following properties are satisfied:*

1. *For each $\theta\in\boldsymbol{P}$ the system $(A(\theta),B(\theta))$ is reachable.*
2. *For finitely many parameters $\theta_1,\ldots,\theta_s\in\boldsymbol{P}$, the $m\times m$ polynomial matrices $D_{\theta_1}(z),\ldots,D_{\theta_s}(z)$ are mutually left coprime.*
3. *For $m+1$ distinct parameters $\theta_1,\ldots,\theta_{m+1}\in\boldsymbol{P}$ the spectra of $A(\theta)$ satisfy*

$$\sigma(A(\theta_1))\cap\cdots\cap\sigma(A(\theta_{m+1}))=\emptyset.$$

4. *Assume $m=1$. The dimension of $\boldsymbol{P}$ satisfies $\dim \boldsymbol{P}\leq 2$. If $A(\theta)$ has a simple real eigenvalue for some $\theta\in\boldsymbol{P}$, then $\dim \boldsymbol{P}\leq 1$.*

Proof. Consider a parameter value $\theta\in\mathbf{P}$ and state vector $\xi\in\mathbb{R}^n$. Choose a continuous map $x^*:\mathbf{P}\longrightarrow\mathbb{R}^n$, with $x^*(\theta)=\xi$. For $\varepsilon>0$ there exists by assumption an input function $u:[0,T]\longrightarrow\mathbb{R}^m$ such that

$$\sup_{\theta\in\mathbf{P}}\|x(T,\theta)-x_*(\theta)\|<\varepsilon.$$

In particular, we obtain $\|x(T,\theta)-\xi\|<\varepsilon$. Thus ξ is in the closure of the reachable set of 0; since the reachable sets of linear systems are closed in $\mathbb{R}^n$, this shows that $(A(\theta),B(\theta))$ is reachable. By the same reasoning, the ensemble reachability of the family $(A(\theta),B(\theta))_\theta$ implies reachability for the parallel interconnection

$$\bar{A}:=\begin{pmatrix}A(\theta_1)&&0\\&\ddots&\\0&&A(\theta_s)\end{pmatrix},\quad \bar{B}:=\begin{pmatrix}B(\theta_1)\\\vdots\\B(\theta_s)\end{pmatrix}\tag{12.7}$$

of finitely many linear systems $(A(\theta_i),B(\theta_i)),i=1,\ldots,s$. By Theorem 10.2, the parallel interconnection (12.7) of reachable linear systems is reachable if and only if the $m\times m$ polynomial matrices $D_{\theta_1}(z),\ldots,D_{\theta_s}(z)$ are mutually left coprime. This completes the proof of the second claim.

The reachability of (12.7) implies that there are at most m Jordan blocks in $\bar{A}$ for each eigenvalue of $\bar{A}$. Thus $\sigma(A(\theta_1))\cap\cdots\cap\sigma(A(\theta_s))=\emptyset$ is satisfied for $s\geq m+1$ distinct parameters $\theta_1,\ldots,\theta_s$, because otherwise there would exist an eigenvalue of $\bar{A}$ with at least $m+1$ Jordan blocks. This proves the third claim.

The last claim follows from the third claim. In fact, let $\lambda(\theta)$ denote a branch of the eigenvalues of $A(\theta)$. Since the eigenvalues of a matrix depend continuously on the parameters θ, one concludes from the fourth claim that the functions $\theta\mapsto\lambda(\theta)\in\mathbb{C}=\mathbb{R}^2$ are continuous and injective. Therefore, since continuous injective functions do not increase dimensions, one concludes that $\dim \mathbf{P}\leq\dim \mathbb{R}^2=2$. Moreover, if there exists a real branch of eigenvalues $\lambda(\theta)$ of $A(\theta)$, then $\dim \mathbf{P}\leq 1$.

If $\lambda(\theta_0)$ is a simple eigenvalue of $A(\theta_0)$, then there exists an open neighborhood U of θ_0 in $\mathbf{P}$ such that for all $\theta \in U$ the eigenvalue $\lambda(\theta)$ is real. This completes the proof. ■

The preceding proof, using Theorem 10.2, shows an interesting connection between ensemble reachability for *finite* parameter sets $\mathbf{P}$ and reachability for parallel interconnection schemes of single-input systems.

Corollary 12.4. *Assume that $\mathbf{P} = \{\theta_1, \ldots, \theta_s\} \subset \mathbb{R}^d$ is finite. Then a family of single-input systems $(A(\theta), b(\theta))_{\theta \in \mathbf{P}}$ is uniformly ensemble reachable if and only if the following two conditions are satisfied:*

1. *$(A(\theta_i), b(\theta_i))$ is reachable for $i = 1, \ldots, s$.*
2. *The characteristic polynomials $\det(zI - A(\theta_i))$ and $\det(zI - A(\theta_j))$ are coprime for all $i \neq j$.*

Proof. This is an obvious consequence of Proposition 12.3. ■

In the discrete-time case, as is further explained in Chapter 10.1, one can strengthen this result by deriving explicit formulas for the inputs that steer to a desired state. In fact, the minimum-time ensemble control task for finite parameter sets becomes equivalent to the Chinese remainder theorem. We illustrate this approach for single-input systems. Let $\theta_1, \ldots, \theta_s \in \mathbf{P}$. The uniform ensemble reachability of the finite family $(A(\theta_i), b(\theta_i))$ is equivalent to the systems $(A_i, b_i) := (A(\theta_i), b(\theta_i)), i = 1, \ldots, s$ being reachable, with pairwise coprime characteristic polynomials $q_i(z) = \det(zI - A_i)$. Define $\hat{q}_i(z) := \prod_{j \neq i} q_j(z)$ and $q(z) := \prod_{j=1}^{n_s} q_j(z)$. Without loss of generality, we can assume that (A_i, b_i) are in controllability canonical form with local state spaces

$$X_{q_j} := \{p \in \mathbb{R}[z] \mid \deg p < \deg q_j = n_j\}.$$

Consider the parallel connection system

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix}.$$

The state space of this global system is

$$X_q := \{p \in \mathbb{R}[z] \mid \deg p < \deg q = n\},$$

with direct sum decomposition

$$X_q := \hat{q}_1(z) X_{q_1} \oplus \cdots \oplus \hat{q}_s(z) X_{q_s}.$$

Assume that local target state vectors $x_j^* \in \mathbb{R}^{n_j}$ are chosen. These vectors uniquely define polynomial elements $r_j(z) \in X_{q_j}$ of degree $< n_j$ via

$$x_j^* = (x_{0j}^*, \dots, x_{n_j-1,j}^*) \quad \text{and} \quad r_j(z) = \sum_{i=0}^{n_j-1} x_{ij}^* z^i.$$

Thus the components of x_j^* are simply the coefficients of the polynomial r_j. The ensemble control goal is then to find a polynomial $f(z) \in X_q$ such that its remainder modulo q_j is r_j. In fact, the coefficients $u_0, \dots, u_{n-1}$ of the polynomial $f(z) = \sum_{i=0}^{n-1} u_{n-i-1} z^i$ are then simply the desired inputs that steer the system from zero to the local states x_j^*. In particular, the minimum length of such an ensemble control is $n = \deg q = \sum_{j=1}^{s} n_j$, as it should be. To compute $f(z)$, we apply the Bezout identity. Thus, by the coprimeness of $q_j, \hat{q}_j$, there exist unique polynomials $a_j(z)$ of degree $< n_j$ and $b_j(z)$ with

$$a_j(z)\hat{q}_j(z) + b_j(z)q_j(z) = 1. \tag{12.8}$$

Define

$$f(z) = \sum_{j=1}^{s} r_j(z)a_j(z)\hat{q}_j(z). \tag{12.9}$$

The Chinese remainder theorem then asserts that f is a unique polynomial of degree n that has r_j as remainder modulo q_j. The coefficients of f thus give the desired controls for (A,b).

Example 12.5. In the discrete-time case and for the parallel connection of s harmonic oscillators, it is very easy to carry out the calculations. For $\theta_1 < \cdots < \theta_s$, let

$$A_j := \begin{pmatrix} 0 & -\theta_j^2 \\ 1 & 0 \end{pmatrix}, \quad b_j := \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Thus (A_j, b_j) is in controllability canonical form, with $q_j(z) = z^2 + \theta_j^2, j = 1, \dots, s$, pairwise coprime. In this example, the Bezout equation (12.8) is easily solved by the constant polynomial

$$a_k(z) = \prod_{j \neq k} (\theta_j^2 - \theta_k^2)^{-1}, \quad k = 1, \dots, s.$$

For the local states $r_j(z) = \xi_1(\theta_j) + \xi_2(\theta_j)z$, formula (12.9) for the remainder polynomial is then

$$f(z) = \sum_{j=1}^{s} (\xi_1(\theta_j) + z\xi_2(\theta_j)) \prod_{k \neq j} \frac{z^2 + \theta_k{}^2}{\theta_k^2 - \theta_j^2}.$$

Note that this is exactly the degree $2s-1$ **Lagrange interpolation polynomial** that satisfies

$$f(\pm\sqrt{-1}\theta_j) = r_j(\pm\sqrt{-1}\theta_j)$$

for $j = 1, \ldots, s$. However, for equidistant choices of the interpolation points, this solution suffers from the well-known Runge phenomenon. Thus, while $f(z)$ is a perfect match to the data at $z = \sqrt{-1}\theta_j$, the approximation error will blow up at the boundary points of the interval **P**. Better approximants can be obtained by interpolating at Chebyshev points.

It is considerably harder to establish sufficient conditions for uniform ensemble control, even if one restricts oneself to single-input systems depending on a scalar parameter θ. Li (2011) has proposed an operator-theoretic characterization of L^2-ensemble reachability for general time-varying linear multivariable systems. However, that characterization is stated in terms of the growth rates of singular values of the input-state operator and, thus, is difficult to verify, even for the time-invariant linear systems (12.1). We next state a result that leads to easily verifiable conditions.

Theorem 12.6 (Sufficient Condition). *Let $\boldsymbol{P} = [\theta_-, \theta_+]$ be a compact interval. A continuous family $(A(\theta), b(\theta))$ of linear single-input systems is uniformly ensemble reachable (or, more generally, L^q-ensemble reachable for $1 \leq q \leq \infty$) provided the following conditions are satisfied:*

(a) $(A(\theta), b(\theta))$ is reachable for all $\theta \in \boldsymbol{P}$.
(b) For pairs of distinct parameters $\theta, \theta' \in \boldsymbol{P}, \theta \neq \theta'$, the spectra of $A(\theta)$ and $A(\theta')$ are disjoint:

$$\sigma(A(\theta)) \cap \sigma(A(\theta')) = \emptyset.$$

(c) For each $\theta \in \boldsymbol{P}$ the eigenvalues of $A(\theta)$ have an algebraic multiplicity of one.

Conditions (a) and (b) are also necessary for uniform ensemble reachability.

Static output feedback control presents an interesting situation where all these assumptions fall easily into place. Thus, consider a fixed reachable and observable linear system (A, b, c). Let $\mathbf{P} = [\theta_-, \theta_+]$ denote a compact interval of gain parameters. Then, for real values of θ, the closed-loop characteristic polynomial is $\det(A - \theta bc) = q(z) + \theta p(z)$, with p, q coprime and

$$c(zI - A)^{-1} b = \frac{p(z)}{q(z)}.$$

In particular, for distinct numbers $\theta \neq \theta'$ there exists no complex number z with $q(z) + \theta p(z) = 0 = q(z) + \theta' p(z)$. Theorem 12.6 therefore implies the following corollary.

Corollary 12.7. *Let (A, b, c) be a reachable and observable SISO system. The (discrete-time or continuous-time) output feedback system $(A - \theta bc, b)$ is uniformly ensemble reachable if the eigenvalues of $A - \theta bc$ are simple for all $\theta \in [\theta_-, \theta_+]$.*

■

The proof of Theorem 12.6 is much easier for discrete-time systems, and therefore we first prove the result for discrete-time single-input systems. In this situation the uniform ensemble reachability condition can be restated in a more convenient form.

Lemma 12.8. *A family $\{(A(\theta), b(\theta)), \theta \in \boldsymbol{P}\}$ of discrete-time single-input systems is uniformly ensemble reachable on $\boldsymbol{P}$ if, for all $\varepsilon > 0$ and all continuous functions $x^* : \boldsymbol{P} \to \mathbb{R}^n$, there exists a real scalar polynomial $p(z) \in \mathbb{R}[z]$ such that*

$$\sup_{\theta \in \boldsymbol{P}} \| p(A(\theta))\, b(\theta) - x^*(\theta) \| < \varepsilon.$$

Proof. For an input sequence $u(0), \ldots, u(T-1)$ the solution is

$$x(T, \theta) = \sum_{k=0}^{T-1} A(\theta)^{kb}(\theta) u(T-1-k) = p(A(\theta))b(\theta),$$

where $p(z) = \sum_{k=0}^{T-1} u_{T-k-1} z^k$. Thus the input sequence is parameter independent if and only if the polynomial $p(z)$ is parameter independent. ■

Using this basic observation we can characterize the uniform ensemble reachability property in explicit form as follows.

Proposition 12.9. *Assume that the discrete-time system $(A(\theta), b(\theta))$ is reachable for $\theta \in \boldsymbol{P} = [\theta_-, \theta_+]$. The following statements are equivalent:*

(a) $(A(\theta), b(\theta))_\theta$ *is uniformly ensemble reachable.*
(b) *For continuous families of polynomials $u_\theta(z) \in \mathbb{R}[z]$ of degree $< n$ and $\varepsilon > 0$, there exists a polynomial $p \in \mathbb{R}[z]$ with*

$$\| p(A(\theta))b(\theta) - u_\theta(A(\theta))b(\theta) \| < \varepsilon.$$

(c) *For continuous families of polynomials $u_\theta(z) \in \mathbb{R}[z]$ of degree $< n$ and $\varepsilon > 0$, there exists a scalar polynomial $p(z) \in \mathbb{R}[z]$ with $\|(p(A(\theta)) - u_\theta(A(\theta))\| < \varepsilon$.*

Assume that for each $\theta \in \boldsymbol{P}$ the eigenvalues of $A(\theta)$ are distinct. Let

$$C := \{(z, \theta) \in \mathbb{C} \times \boldsymbol{P} \mid \quad \det(zI - A(\theta)) = 0\}.$$

Then each of the preceding conditions is equivalent to the following statements:

(d) For continuous families of polynomials $u_\theta(z) \in \mathbb{R}[z]$ of degree $< n$ and a $\varepsilon > 0$, there exists a polynomial $p \in \mathbb{R}[z]$ with

$$|p(z) - u_\theta(z)| < \varepsilon \quad \forall (z,\theta) \in C.$$

Proof. Let

$$R(A,b) = (b, Ab, \ldots, A^{n-1}b)$$

denote the $n \times n$ reachability matrix. By reachability, the matrix $R(A(\theta), b(\theta))$ is invertible for each $\theta \in \mathbf{P}$. For $x^* : \mathbf{P} \longrightarrow \mathbb{R}^n$ continuous, define a polynomial $u_\theta(z) \in \mathbb{R}[z]$ of degree $< n$ as

$$u_\theta(z) = (1, z, \ldots, z^{n-1}) R(A(\theta), b(\theta))^{-1} x^*(\theta).$$

Conversely, every continuous family of polynomials $u_\theta(z)$ of degree $< n$ can be written in this way. Since $u_\theta(A(\theta))b(\theta) = x^*(\theta)$, Lemma 12.8 implies the equivalence of (a) with (b). Obviously, condition (c) implies (b). Assume that the estimate $\|(f - u_\theta)(A(\theta))b(\theta)\| < \varepsilon$ holds. Then $\|(f - u_\theta)(A(\theta))A(\theta)^{kb}(\theta)\| < \varepsilon \cdot \sup_{\theta \in \mathbf{P}} \|A(\theta)\|^k$. Therefore,

$$\|(f - u_\theta)(A(\theta))\| < c\varepsilon$$

for the constant $c = \sup_{\theta \in \mathbf{P}} \|R(A(\theta), b(\theta))^{-1}\| \max_{0 \le k \le n-1} \|A(\theta)\|^k$. Thus (b) implies (c). Now consider a matrix X with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Then, for real polynomials F and $\varepsilon > 0$, the operator norm bound $\|F(X)\| < \varepsilon$ is equivalent to $|F(\lambda_i)| < \delta(\varepsilon)$, $i = 1, \ldots, n$. Here $\delta(\varepsilon)$ goes to zero if and only if ε goes to zero. This shows that condition (d) is equivalent to (c), and we are done. ■

1. Proof of Theorem 12.6 for Discrete-Time Systems. We now prove Theorem 12.6 for discrete-time systems. Our proof depends on Mergelyan's theorem, a rather deep theorem from complex approximation theory. It is stated here for convenience; see Chapter II in Gaier (1987) for a proof. Let $\hat{\mathbb{C}}$ denote the one-point compactification of $\mathbb{C}$, i.e., the complex plane which is extended by including the point ∞.

Theorem 12.10 (Mergelyan). *Suppose K is compact in $\hat{\mathbb{C}}$ such that the complement $\hat{\mathbb{C}} \setminus K$ is connected. Suppose that $f : K \longrightarrow \mathbb{C}$ is a continuous function that is holomorphic in the interior of K. Then for every $\varepsilon > 0$ there exists a complex polynomial $p(z) \in \mathbb{C}[x]$ such that, for all $z \in K$,*

$$|f(z) - p(z)| < \varepsilon.$$

We note that this theorem applies in particular in the special case where K has no interior points. In that case, the analyticity condition on f is superfluous.

Proof. The claims in Theorem 12.6 concerning necessity all follow from Lemma 12.3. Consider the compact set

$$C := \{(z,\theta) \in \mathbb{C} \times \mathbf{P} | \quad \det(zI - A(\theta)) = 0\}$$

and the projection map $\pi : C \longrightarrow \mathbb{C}$ defined by $\pi(z,\theta) = z$. Since $\mathbf{P}$ is compact, it follows that C and, therefore, the image $K := \pi(C)$ are compact. Condition (b) of Theorem 12.6 is equivalent to π being injective on C, and therefore $\pi : C \longrightarrow K$ is a homeomorphism onto K. Thus, there exists a continuous map $\sigma : K \longrightarrow C, \sigma(z) = (z, \tau(z))$ that is a right inverse of π. The continuous map $\tau : K \longrightarrow \mathbf{P}$ has the property that $z \in K \iff (z, \tau(z)) \in C$. Since the eigenvalues of $A(\theta)$ define the continuous functions of θ and are assumed to be distinct for each $\theta \in \mathbf{P}$, one concludes that $K = K_1 \cup \cdots \cup K_n \subset \mathbb{C}$ consists of exactly n disjoint homeomorphic copies K_j of $\mathbf{P}$. Since $\mathbf{P}$ is simply connected, so is K, and thus the complement $\mathbb{C} \setminus K$ is connected. Moreover, K has an empty interior. Therefore, one can apply Mergelyan's Theorem 12.10 to uniformly approximate continuous functions F on K by polynomials. For a continuous family of desired states $x^*(\theta)$, define the polynomial u_θ in z as

$$u_\theta(z) = (1, z, \ldots, z^{n-1}) R(A(\theta), b(\theta))^{-1} x^*(\theta).$$

Replacing the variable θ in u_θ with $\theta = \tau(z)$ we obtain a continuous function

$$F : K \longrightarrow \mathbb{C}, \quad F(z) := u_{\tau(z)}(z).$$

Thus the theorem by Mergelyan asserts that there exists a polynomial $p(z)$, with $|p(z) - F(z)| < \varepsilon$, uniformly on K. Equivalently, there exists a polynomial $p(z)$ such that

$$|p(z) - u_\theta(z)| < \varepsilon \quad \forall (z,\theta) \in C. \tag{12.10}$$

Note that $u_\theta(z)$ is a real polynomial. Thus, by a possible replacement of $p(z)$ by $\frac{1}{2}(p(z) + \overline{p}(z))$, we can assume that (12.10) is satisfied for a real polynomial $p(z)$. The result follows from Proposition 12.9. This completes the proof of Theorem 12.6 in the discrete-time case. ■

2. Proof of Theorem 12.6 for Continuous-Time Systems. For continuous-time systems

$$\dot{x} = A(\theta)x(t,\theta) + b(\theta)u(t), \quad x_0(\theta) = 0,$$

we apply a sampling argument. For each positive sampling period $\tau > 0$, consider the discrete-time system

$$x(t+1,\theta) = F(\theta)x(t,\theta) + g(\theta)u(t), \quad x_0(\theta) = 0, \tag{12.11}$$

where

$$F(\theta) := e^{\tau A(\theta)}, \quad g(\theta) = \left(\int_0^\tau e^{sA(\theta)} ds\right) b(\theta).$$

The result now follows from showing the next proposition.

Proposition 12.11. *Let $(A(\theta), b(\theta))$ satisfy the assumptions of Theorem 12.6. Then the sampled system (12.11) satisfies the assumptions, too.*

Proof. By the compactness of $\mathbf{P}$, there exists $\tau_* > 0$ such that all conditions of Theorem 12.6 are satisfied for the discrete-time system $(F(\theta), G(\theta))$ and all $0 < \tau < \tau_*$. It is well known that the reachability of a continuous-time linear system (A, b) implies the reachability of the sampled discrete-time system (F, g) if the sampling period is sufficiently small. Thus the pointwise reachability condition (a) implies the same condition for $(F(\theta), g(\theta))$. The other conditions follow from the continuity of the eigenvalues and local injectivity of the matrix exponential function. ■

Applying Proposition 12.11, the proof of Theorem 12.6 for the discrete-time case implies the uniform ensemble reachability of the discrete-time system (12.11). Note that, under sampling, the continuous- and discrete-time solutions coincide at the sampling points. Therefore, the finite-length input sequence u_k for the uniform ensemble reachability of $(F(\theta), g(\theta))$ induces a piecewise constant input function $u^\tau : [0, T] \longrightarrow \mathbb{R}$ that performs the uniform ensemble control task for the continuous-time system (12.1). This completes the proof of our main theorem. ■

As mentioned earlier, conditions (a) and (b) in Theorem 12.6 are actually necessary conditions. The next result shows that the ensemble reachability of discrete-time systems can be shown under weaker assumptions than condition (c).

Proposition 12.12. *Let $\mathbf{P} = [\theta_-, \theta_+]$ be a compact interval. A continuous family $(A(\theta), b(\theta))$ of linear discrete-time single-input systems is uniformly ensemble reachable provided the following conditions are satisfied:*

(a) *$(A(\theta), b(\theta))$ is reachable for all $\theta \in \mathbf{P}$.*
(b) *For pairs of distinct parameters $\theta, \theta' \in \mathbf{P}, \theta \neq \theta'$, the spectra of $A(\theta)$ and $A(\theta')$ are disjoint:*

$$\sigma(A(\theta)) \cap \sigma(A(\theta')) = \emptyset.$$

(c) *The characteristic polynomial of $A(\theta)$ is of the form $z^n - a_{n-1}z^{n-1} - \cdots - a_1 z - a_0(\theta)$, with $a_1, \ldots, a_{n-1}$ constant real numbers.*

Proof. Without loss of generality, we can assume that $A(\theta), b(\theta)$ is in controllability normal form for each θ. The polynomial $\pi(z) := z^n - a_{n-1}z^{n-1} - \cdots - a_1 z$ satisfies $\pi(A(\theta)) = a_0(\theta) I_n$. Moreover, $A(\theta)^{k-1} e_1 = e_k$ for $k = 1, \ldots, n$. By condition (b), we see that $a_0 : \mathbf{P} \longrightarrow \mathbb{R}$ is injective, and hence the inverse function

$a_0^{-1} : a_0(\mathbf{P}) \longrightarrow \mathbf{P}$ exists and is continuous. Using the Weierstrass approximation theorem there exist polynomials $f_1(t), \dots, f_n(t) \in \mathbb{R}[t]$ such that, for $k = 1, \dots, n$,

$$\sup_{t \in a_0(\mathbf{P})} |f_k(t) - e_k^\top x_*(a_0^{-1}(t)| < \varepsilon$$

or, equivalently, $\sup_{\theta \in \mathbf{P}} |f_k(a_0(\theta)) - e_k^\top x_*(\theta)| < \varepsilon$. The real polynomial $f(z) := \sum_{k=1}^n f_k(\pi(z)) z^{k-1}$ satisfies

$$f(A(\theta))e_1 = \sum_{k=1}^n f_k(\pi(A(\theta)))A(\theta)^{k-1}e_1 = \sum_{k=1}^n f_k(a_0(\theta))e_k.$$

This implies $\sup_{\theta \in \mathbf{P}} \|f(A(\theta))b(\theta) - x_*(\theta)\|_\infty < \varepsilon$, and the result follows. ■

We illustrate the applicability of the preceding results by means of three examples.

Example 12.13 (L^2-Ensemble Observability). We briefly discuss the dual version of ensemble reachability, i.e., ensemble observability. Since duality theory is easier in a Hilbert-space context, we focus on the notions of L^2-ensemble observability.

Definition 12.14. Assume that $A(\theta) \in \mathbb{R}^{n \times n}, C(\theta) \in \mathbb{R}^{p \times n}$ vary continuously in a compact parameter domain $\mathbf{P} \subset \mathbb{R}^d$. The parameter-dependent system

$$\begin{aligned} \frac{\partial x(t,\theta)}{\partial t} &= A(\theta)x(t,\theta), \quad x(0,\cdot) \in L^2(\mathbf{P}, \mathbb{R}^n), \\ y(t) &= \int_{\mathbf{P}} C(\theta)x(t,\theta)d\theta \end{aligned} \tag{12.12}$$

is called L^2**-ensemble observable** if there exists $T > 0$ such that $y(t) = 0$ on $[0,T]$ implies $x(0,\theta) = 0$ for all $\theta \in \mathbf{P}$.

Definition 12.14 implies that one can reconstruct the L^2-initial state $x(0,\cdot)$ of (12.12) from the average values

$$\int_{\mathbf{P}} C(\theta)x(t,\theta)d\theta, \quad 0 \le t \le T,$$

of the outputs $C(\theta)x(t,\theta)$. Thus ensemble observability is a rather strong property that is particularly useful in, for example, biological parameter identification tasks where often only an averaged type of output information is available.

System (12.12) is equivalent to the linear system

$$\begin{aligned} \dot{x}(t) &= \mathscr{A}x(t), \quad x(0) \in L^2(\mathbf{P}, \mathbb{R}^n), \\ y(t) &= \mathscr{C}x(t) \end{aligned} \tag{12.13}$$

on the Hilbert space $X = L^2(\mathbf{P}, \mathbb{R}^n)$. Here $\mathscr{A} : X \longrightarrow X, \mathscr{C} : X \longrightarrow \mathbb{R}^p$ are bounded linear operators defined by

$$(\mathscr{A}x)(\theta) = A(\theta)x(\theta), \quad (\mathscr{C}x)(\theta) = \int_{\mathbf{P}} C(\theta)x(\theta)d\theta,$$

respectively. Thus, $\mathscr{A}$ is a multiplication operator while $\mathscr{C}$ is an integration operator. The preceding notion of ensemble observability is equivalent to the notion of the approximate observability of (12.13), as defined in Curtain and Zwart (1995), Definition 4.1.12. Moreover, Lemma 4.1.13 in Curtain and Zwart (1995) implies that (12.13) is approximate observable if and only if the dual system

$$\frac{\partial}{\partial t}x(t,\theta) = A(\theta)^\top x(t,\theta) + C(\theta)^\top u(t), \quad x(0,\theta) = 0, \tag{12.14}$$

is L^2-ensemble reachable. Therefore, Theorem 12.6 applies to (12.14) for $p = 1$. This shows that every continuous one-parameter family $(A(\theta), C(\theta)), \theta \in \mathbf{P} = [\theta_-, \theta_+]$, of single-output linear systems is L^2-ensemble observable provided the following three conditions are satisfied:

1. $(A(\theta), C(\theta))$ is observable for all $\theta \in \mathbf{P}$.
2. The spectra of $A(\cdot)$ are pairwise disjoint, i.e.,

$$\sigma(A(\theta)) \cap \sigma(A(\theta')) = \emptyset, \quad \forall \theta, \theta' \in \mathbf{P}, \theta \neq \theta.$$

3. For each $\theta \in \mathbf{P}$ the eigenvalues of $A(\theta)$ have an algebraic multiplicity of one.

Example 12.15 (Robust Numerical Integration.). The simplest numerical integration method is certainly the Euler-step method. For a continuous-time linear control system $\dot{x} = Ax + Bu$, this yields the discrete-time system

$$x(t+1,h) = (I + hA)x(t,h) + hBu(t), \tag{12.15}$$

with a step-size parameter $h > 0$. One can then ask whether the family of discretized systems (12.15) can be robustly controlled using a control sequence $u(t), t \in \mathbb{N}$, that is independent of the step-size $h > 0$. Theorem 12.6 provides a simple answer. Assume that the pair (A, B) is reachable. Then for each parameter $h > 0$ the pairs $(I + hA, hB)$ are also reachable. Moreover, assume that A has only simple, distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ that satisfy

$$\lambda_j \neq c\lambda_i \tag{12.16}$$

for all $c > 0$ and $i \neq j$. Then the eigenvalues $1 + h\lambda_1, \ldots, 1 + h\lambda_n$ of $I + hA$ are also simple and satisfy $1 + h\lambda_i \neq 1 + h'\lambda_j$ for $h \neq h'$. Thus the discretized system (12.15) is ensemble reachable for compact intervals $\mathbf{P} \subset (0, \infty)$ of step-size parameters provided (A, B) is reachable and A has distinct simple eigenvalues that satisfy (12.16).

This shows that the Euler-step approximation (12.15) of a continuous-time control system $\dot{x} = Ax + Bu$ inherits the reachability properties in a very strong sense, i.e., (12.15) can be controlled in a step-size independent way. It would be interesting to see whether this property carried over to more general, higher-order, Runge–Kutta methods.

Example 12.16 (Robust Open-Loop Synchronization.). We describe an application of Theorem 12.6 to the synchronization of N identical harmonic oscillators

$$\ddot{y}(t) + \omega^2 y(t) = v(t),$$

with state-space realization

$$A := \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}, \quad b := \begin{pmatrix} \frac{1}{\omega} \\ 0 \end{pmatrix}, \quad c := \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

The identical frequency ω of the harmonic oscillators is assumed to be known. We assume that the oscillators are coupled in a ring with unknown coupling strength κ, which can vary over a compact interval $\mathbf{P} := [\kappa^-, \kappa^+] \subset (0, \infty)$ of positive numbers. Thus the network topology is described by a directed simple cycle graph with N nodes and weighted adjacency matrix κS, with the circulant matrix

$$S := \begin{pmatrix} 0 & 1 & & 0 \\ 0 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ 1 & & 0 & 0 \end{pmatrix}.$$

We assume that the numbering of the harmonic oscillators is such that a single external input is applied to the first harmonic oscillator. Thus the input-to-state interconnection matrix is $e_1 = (1, 0, \ldots, 0)^\top$. The dynamics of the overall network is thus of the form

$$\begin{aligned} \tfrac{\partial}{\partial t} x(t,\kappa) &= \big(I \otimes A + \kappa S \otimes bc\big) x(t,\kappa) + \big(e_1 \otimes b\big) u(t), \\ x(0,\kappa) &= x^0, \end{aligned} \tag{12.17}$$

where $x^0 \in \mathbb{R}^{2N}$ denotes the initial state of the network. Let $\mathbf{e} = (1, \ldots, 1)^\top$, and let $x^* \in \mathbb{R}^2$ denote the desired terminal state to which the harmonic oscillators are supposed to synchronize. The network of harmonic oscillators (12.17) is called **robustly synchronizable** from $x^0 \in \mathbb{R}^{2N}$ to $\mathbf{e} \otimes x^* \in \mathbb{R}^{2N}$ if for every $\varepsilon > 0$ there exists $T > 0$ and an input function $u \colon [0, T] \to \mathbb{R}$ such that the state $x(t)$ of (12.17) satisfies

$$\sup_{\kappa \in \mathbf{P}} \| x(T,\kappa) - \mathbf{e} \otimes x^* \| < \varepsilon.$$

We emphasize that the input u acts here as a universal input for the parameter-dependent network that steers x^0 to the synchronized state $\mathbf{1}\otimes x^*$ uniformly for all coupling strength parameters $\kappa \in \mathbf{P}$.

We next show, using Theorem 12.6, that the network (12.17) of harmonic oscillators robustly synchronizes from every initial state $x^0 \in \mathbb{R}^{2N}$ at a vector $\mathbf{1}\otimes x^* \in \mathbb{R}^{2N}$ of identical terminal states x^*. Let $\boldsymbol{\Phi}$ denote the $N\times N$ Fourier matrix (9.39) and $\omega = e^{2\pi\sqrt{-1}/N}$. Thus, using (9.40), it follows that $\boldsymbol{\Phi}^* S\boldsymbol{\Phi} = \operatorname{diag}\left(1,\omega,\ldots,\omega^{N-1}\right)$ and $\boldsymbol{\Phi} e_1 = \frac{1}{\sqrt{N}}\mathbf{e}$. Applying the state-space similarity transformation $z := (\boldsymbol{\Phi}^*\otimes I_n)x$ using the unitary matrix $\boldsymbol{\Phi}^*\otimes I_n$, we see that (12.17) is similar to the decoupled, parameter-dependent system

$$\begin{aligned}\frac{\partial}{\partial t} z_j(t,\kappa) &= \left(A+\kappa e^{2\pi j\sqrt{-1}/N})bc\right) z_j(t,\kappa)+\sqrt{N}bu(t), \quad j=1,\ldots,N\\ z(0,\kappa) &= \boldsymbol{\Phi}^* x^0.\end{aligned}$$

This system is the parallel connection of reachable linear systems. Moreover, the eigenvalues of $I\otimes A+\kappa S\otimes bc$ are equal to

$$\bigcup_{j=1}^{N}\left\{z\in\mathbb{C} : z^2+\omega^2-\kappa\omega^j=0\right\}.$$

These eigenvalues are distinct and simple if, for example, $1\notin \mathbf{P} = [\kappa^-,\kappa^+]$. Moreover, under this condition, the eigenvalues for different $\kappa\neq\kappa'$ are distinct. This implies that $I\otimes A+\kappa S\otimes bc$ is reachable and conditions (a)–(c) of Theorem 12.6 are satisfied. One concludes that the network (12.17) is robustly synchronizable from every initial state provided $1\notin[\kappa^-,\kappa^+]$. Similarly, robust synchronization can be established for more complicated network structures.

12.3 Control of Platoons

The Fourier transform provides an elegant way to utilize the preceding results on uniform ensemble control for the control of partial differential equations and platoons of systems. Although this point of view has appeared already in previous work by, for example, Green and Kamen (1985), Bamieh, Paganini and Dahleh (2002), and Curtain, Iftime and Zwart (2009), the control tasks that we consider here have not been addressed before. In fact, the previously cited works consider only scenarios where each subsystem is controlled by individual, independent input functions. Thus such approaches use an infinite number of control functions and are therefore severely limited in applicability. In contrast, we focus on the reachability of systems that employ a finite number of input functions that are distributed over the entire network. For controlling platoons or swarms of systems this approach appears to be more appropriate.

1. Finite Platoons Consider the task of controlling a finite platoon of N vehicles on a line. If each vehicle is controlled individually by independent input functions, then clearly the overall system will be reachable. A more interesting situation arises when one considers inputs that are broadcasted to all systems simultaneously. Thus all vehicles are controlled by the same input function (or by a small number of input functions). By assuming nearest-neighbor interactions, we obtain the control system

$$\begin{aligned}
\dot{x}_1 &= -x_2 + x_1 + u(t)\\
\dot{x}_2 &= -x_3 + 2x_2 - x_1 + u(t)\\
&\vdots\\
\dot{x}_{N-1} &= -x_N + 2x_{N-1} - x_{N-2} + u(t)\\
\dot{x}_N &= -x_{N-1} + x_N + u(t).
\end{aligned}$$

In matrix form the system is $\dot{x}(t) = Ax(t) + bu(t)$, where

$$A = \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & \ddots & \ddots & & \\ & & \ddots & \ddots & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix}. \tag{12.18}$$

To check for reachability, we apply the Hautus test. Recall from Theorem 8.46 that the eigenvalues of A are nonrepeated and are equal to $\lambda_k = 2 - 2\cos\frac{(k-1)\pi}{N}, k = 1, \ldots, N$. Moreover, $\xi^{(k)} = (\sin\frac{k\pi}{N+1}, \ldots, \sin\frac{kN\pi}{N+1})^\top$ is an eigenvector for the eigenvalue λ_k. Define $\omega = e^{\frac{\sqrt{-1}\pi}{N+1}}$. Thus

$$\begin{aligned}
b^\top \xi^{(k)} &= \sum_{\nu=1}^{N} \sin\frac{k\nu\pi}{N+1} = \mathrm{Im}\sum_{\nu=1}^{N} \omega^{k\nu}\\
&= \mathrm{Im}\left(\frac{1-\omega^{k(N+1)}}{1-\omega^k} - 1\right) = \mathrm{Im}\left(\frac{1-(-1)^k}{1-\omega^k} - 1\right)\\
&= \begin{cases} 0 & \text{if} \quad k \quad \text{is even,}\\ \frac{\sin\frac{k\pi}{N+1}}{1-\cos\frac{k\pi}{N+1}} & \text{if} \quad k \quad \text{is odd} \end{cases}
\end{aligned}$$

is zero if and only if $1 \le k \le N$ is even. This implies that the system is not reachable. In contrast, consider the case where $b = e_k$ for some $1 \le k \le N$. From the tridiagonal structure of A it is easily seen that (A, e_1) is reachable. The subsequent characterization of reachability shows an interesting connection to elementary number theory.

Theorem 12.17. *Let (A, b) be defined by (12.18).*

1. *The pair (A, b) is not reachable.*
2. *(A, e_k) is reachable if and only if k and $N+1$ are coprime.*
3. *$(A, e_k + e_\ell)$ is reachable provided both pairs $k+\ell, N+1$ and $k-\ell, N+1$ are coprime.*
4. *The two-input system $(A, (e_k, e_\ell))$ is reachable if and only if there exists no $1 \leq r \leq N$ such that $N+1$ divides both rk and $r\ell$.*

Proof. The first part has already been shown. For the second part note that $e_k^\top \xi^{(r)} = \sin \frac{kr\pi}{N+1} = 0$ if and only if $N+1$ divides kr. Suppose that k and $N+1$ are coprime. Then $N+1$ divides kr if and only if $N+1$ divides k. But this is impossible because $1 \leq k \leq N$. Thus the coprimeness of k and $N+1$ implies $e_k^\top \xi^{(r)} \neq 0$ for all $1 \leq r \leq N$, i.e., it implies the reachability of (A, e_k). Conversely, assume that $d \geq 2$ is the greatest common divisor of $k, N+1$. Then $k = k'd$ and $N+1 = N'd$ for suitable integers $1 \leq k', N' \leq N$. Then $(N+1)k' = kr$ for $r := N' \leq N$. Thus $e_k^\top \xi^{(r)} = 0$, and therefore (A, e_k) cannot be reachable.

By the Hautus criterion, the reachability of $(A, e_k + e_\ell)$ is violated if and only if there exists $1 \leq r \leq N$ such that $\sin \frac{kr\pi}{N+1} + \sin \frac{\ell r\pi}{N+1} = 0$. Recall that $\sin(\pi x) = \sin(\pi y)$ if and only if either $x+y$ is an odd integer or $x-y$ is an even integer. Thus, reachability holds if and only if $r(k-\ell) \notin (2\mathbb{Z}+1)(N+1)$ and $r'(k+\ell) \notin 2\mathbb{Z}(N+1)$ is valid for all $1 \leq r, r' \leq N$. Certainly this is the case if $k+\ell, N+1$ are coprime and $k-\ell, N+1$ are coprime. This implies the result. Finally, the reachability of the two-input system $(A, (e_k, e_\ell))$ is equivalent to the condition that there exists no $1 \leq r \leq N$ with $\sin \frac{kr\pi}{N+1} = \sin \frac{\ell r\pi}{N+1} = 0$. This proves the result. ∎

Similarly, let us consider the situation where vehicles proceed on a circular domain. In this case, we obtain the linear control system $\dot{x}(t) = Fx(t) + gu(t)$ with system matrices

$$F = \begin{pmatrix} 2 & -1 & & & & -1 \\ -1 & 2 & -1 & & & \\ & -1 & \ddots & \ddots & & \\ & & \ddots & \ddots & -1 & \\ & & & -1 & 2 & -1 \\ -1 & & & & -1 & 2 \end{pmatrix}, \quad g = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix}. \tag{12.19}$$

See Notarstefano and Parlangeli (2013) and Chapter 9 for reachability results on closely related linear systems evolving on circular arrays. The eigenvalues of the circulant matrix F are equal to $2 - 2\cos \frac{2k\pi}{N}, k = 1, \ldots, N$. Thus, for $N \geq 2$, the symmetric matrix F always has eigenvalues with a multiplicity of 2. Therefore, the circulant system (F, g) is not reachable. This also follows directly from the fact that

$Fg = 0$. However, contrary to the preceding case of platoons on a line, the occurrence of eigenvalues with multiplicities ≥ 2 in the Jordan canonical form of F shows that the system (F, g) is not reachable *for each choice* of g. Thus we need at least two independent inputs to control a platoon on a circle. The next result describes some of the possibilities for controlling with two inputs.

Theorem 12.18. *Let $1 \leq k < \ell \leq N$ and F be defined by (12.19). There exists no vector g such that (F, g) is reachable. The two-input system $(F, (e_k, e_\ell))$ is reachable if and only if N and $k - \ell$ are coprime.*

Proof. By Theorem 8.48, the eigenspaces of F are either one-dimensional and spanned by $u := e_1 + \cdots + e_N$ (for $r = N$) or by $u := e_1 - e_2 + e_3 - \cdots + e_N$ (for $N = 2m, r = m$) or are two-dimensional with a basis

$$x^{(r)} = \begin{pmatrix} 1 \\ \cos(\frac{2r\pi}{N}) \\ \cos(\frac{4r\pi}{N}) \\ \vdots \\ \cos(\frac{2(N-1)r\pi}{N}) \end{pmatrix}, \quad y^{(r)} = \begin{pmatrix} 1 \\ \sin(\frac{2r\pi}{N}) \\ \sin(\frac{4r\pi}{N}) \\ \vdots \\ \sin(\frac{2(N-1)r\pi}{N}) \end{pmatrix}.$$

Thus reachability is satisfied whenever the inner products of e_k, e_ℓ with the basis of eigenvectors does not vanish. Computing the inner products we obtain

$$\begin{gathered} < e_k, u > = < e_\ell, u > = 1, \quad < e_k, v > = (-1)^k, \quad < e_\ell, v > = (-1)^\ell, \\ < e_k, x^{(r)} > = \sin \frac{2(k-1)r\pi}{N}, \quad < e_\ell, x^{(r)} > = \sin \frac{2(\ell-1)r\pi}{N}, \\ < e_k, y^{(r)} > = \cos \frac{2(k-1)r\pi}{N}, \quad < e_\ell, x^{(r)} > = \cos \frac{2(\ell-1)r\pi}{N}. \end{gathered}$$

Thus the inner products with the eigenvectors are nonzero for the simple eigenvalues λ_r, with $r = N$ or $r = m, N = 2m$. For the other cases assume that $v = \alpha x^{(r)} + \beta y^{(r)}$ is an eigenvector of F for λ_r, with $< v, e_k > = < v, e_\ell > = 0$. Then

$$\begin{aligned} \alpha \sin \frac{2(k-1)r\pi}{N} + \beta \cos \frac{2(k-1)r\pi}{N} &= 0, \\ \alpha \sin \frac{2(\ell-1)r\pi}{N} + \beta \cos \frac{2(\ell-1)r\pi}{N} &= 0. \end{aligned}$$

This has a nonzero solution (α, β) if and only if $\sin \frac{2(k-\ell)r\pi}{N} = 0$. This is equivalent to the condition that N divides $(k - \ell)r$. Since $r < N$, this implies reachability if and only if N and $k - \ell$ are coprime. ■

Using the theory developed in Chapter 9, it is easy to extend the reachability analysis of platoons from first-order scalar systems to higher-order systems.

As an example, consider a homogeneous network of identical higher-order systems of the form

$$
\begin{aligned}
&q(\frac{d}{dt})y_i(t) = p(\frac{d}{dt})v_i(t)\\
&\qquad v_i(t) = y_{i+1}(t) - 2y_i(t) + y_{i-1}(t) + b_i u(t), \quad i = 1,\ldots,N.
\end{aligned}
\tag{12.20}
$$

Here $b = \mathrm{col}\,(b_1,\ldots,b_N) \in \mathbb{R}^N$, and $p(z)$ and $q(z)$ denote coprime real scalar polynomials with $\deg p < \deg q = n$. Defining $Q(z) = q(z)I_N$, $P(z) = p(z)I_N$, and A as in (12.18) we see that (12.20) is equivalent to the homogeneous network

$$
\left(Q(\frac{d}{dt}) - P(\frac{d}{dt})A\right) y(t) = P(\frac{d}{dt})bu(t).
$$

Applying Theorem 9.15, one concludes that the network (12.20) is reachable if and only if (A,b) is reachable. Thus, for $b = e_k$, the network (12.20) is reachable if and only if $N+1$ and k are coprime, independently of the choice of coprime polynomials $p(z), q(z)$. Similarly, reachability results for platoons of higher-order systems with circulant interconnection matrices are obtained.

2. Infinite Platoons. We now turn to an analysis of infinite platoons and their reachability properties. Infinite platoons are infinite-dimensional control systems where the spatial variable is constrained to either $\mathbb{N}$ or $\mathbb{Z}$. The coordinates of infinite platoons are therefore defined by either one-sided infinite sequences $(x_k)_{k\in\mathbb{N}_0}$ or bi-infinite sequences $(x_k)_{k\in\mathbb{Z}}$ of elements $x_k \in \mathbb{R}^n$. In either case, we obtain a Hilbert space $\ell^2_+(\mathbb{R}^n)$ or $\ell^2(\mathbb{R}^n)$ of square summable sequences with norms

$$
\|x\|^2 = \sum_{k=0}^{\infty} |x_k|^2, \quad \|x\|^2 = \sum_{k=-\infty}^{\infty} |x_k|^2,
$$

respectively. The interpretation of, for example, the set of integers $\mathbb{Z}$ with the spatial domain of a platoon is due to the identification of curves $t \mapsto x(t)$ in $\ell^2(\mathbb{R}^n)$ with functions $x : \mathbb{Z}\times\mathbb{R} \longrightarrow \mathbb{R}^n, (k,t) \mapsto x(k,t) := x(t)_k$. In such an interpretation the space $\mathbb{Z}$ corresponds to the spatial variable, while t corresponds to time. We follow the work by Curtain, Iftime and Zwart (2009), who developed an operator-theoretic analysis of infinite platoons in the Hilbert space $\ell^2(\mathbb{R}^n)$ of bi-infinite sequences. We also refer to the more recent work by Feintuch and Francis (2012) for a deeper analysis of stability problems for platoons in the Banach space of ℓ^∞ sequences. Restricting the coordinates to the Hilbert-space context of ℓ^2 sequences enables us to employ Fourier-transform techniques. In a second step, we then connect the reachability analysis of infinite platoons to that for parameter-dependent linear systems.

Restricting the coordinates to the Hilbert-space context of ℓ^2 sequences enables us to employ Fourier-transform techniques. For bi-infinite sequences x of vectors $x_k \in \mathbb{C}^n, k \in \mathbb{Z}$, define the associated Fourier series

$$x(e^{\sqrt{-1}\theta}) := \sum_{k=-\infty}^{\infty} x_k e^{-\sqrt{-1}k\theta}.$$

Conversely, with functions $f \in L^2(S^1, \mathbb{C}^n)$ on the unit circle S^1 one associates the sequence of Fourier coefficients

$$x_k(f) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{\sqrt{-1}\theta}) e^{\sqrt{-1}k\theta} d\theta, \quad k \in \mathbb{Z}.$$

The **discrete Fourier transform**

$$\mathscr{F} : \ell^2(\mathbb{C}^n) \longrightarrow L^2(S^1, \mathbb{C}^n), \quad \mathscr{F}(x) = x(e^{\sqrt{-1}\theta})$$

then defines a linear isomorphism whose inverse is defined by the Fourier coefficients via

$$\mathscr{F}^{-1} : L^2(S^1, \mathbb{C}^n) \longrightarrow \ell^2(\mathbb{C}^n), \quad \mathscr{F}^{-1} f = (x_k(f))_{k \in \mathbb{Z}}.$$

By the Plancherel formula, the discrete Fourier transform defines an isometry of Hilbert spaces, i.e., for all $x \in \ell^2(\mathbb{C}^n)$,

$$\|x\|_{\ell^2} = \|\mathscr{F}x\|_{L^2}.$$

Similarly, the discrete Fourier transform of a one-sided sequence $(x_k)_{k \geq 0} \in \ell^2_+(\mathbb{C}^n)$ is defined as

$$x(e^{\sqrt{-1}\theta}) := \sum_{k=0}^{\infty} x_k e^{-\sqrt{-1}k\theta},$$

which defines an element of the Hardy space $H^2(\mathbb{D}; \mathbb{C}^n)$ on the unit disc $\mathbb{D}$. The discrete Fourier transform

$$\mathscr{F} : \ell^2_+(\mathbb{C}^n) \longrightarrow H^2(\mathbb{D}; \mathbb{C}^n), \quad \mathscr{F}(x) = x(e^{\sqrt{-1}\theta})$$

then maps $\ell^2_+(\mathbb{C}^n)$ isometrically onto $H^2(\mathbb{D}; \mathbb{C}^n) = H^2(\mathbb{D})^n$, the Hardy space of n-tuples of holomorphic functions on the open unit disc that are Lebesgue square integrable on the unit circle.

Following these preliminaries, we now turn to a study of bi-infinite platoons on $\mathbb{Z}$. The simplest classes of such models are spatially invariant and have the form

$$\dot{x}_k(t) = \sum_{j=-\infty}^{\infty} A_j x_{k-j}(t) + \sum_{j=-\infty}^{\infty} B_j u_{k-j}(t), \quad k \in \mathbb{Z}, \tag{12.21}$$

where $A_k \in \mathbb{R}^{n \times n}$ and $B_k \in \mathbb{R}^{n \times m}$. Here we assume that the inputs and state sequences are in ℓ^2, i.e., $(x_i) \in X = \ell^2(\mathbb{C}^n)$ and $(u_i) \in U = \ell^2(\mathbb{C}^m)$, respectively.

To specify conditions that the convolution operators $\mathscr{A} : X \longrightarrow X$ and $\mathscr{B} : U \longrightarrow X$, with

$$(\mathscr{A}x)_k = \sum_{j=-\infty}^{\infty} A_j x_{k-j}, \quad (\mathscr{B}u)_k = \sum_{j=-\infty}^{\infty} B_j u_{k-j},$$

are well defined, we impose a condition on their Fourier symbols

$$A(e^{\sqrt{-1}\theta}) = \sum_{j=-\infty}^{\infty} A_j e^{-\sqrt{-1}j\theta}, \quad B(e^{\sqrt{-1}\theta}) = \sum_{j=-\infty}^{\infty} B_j e^{-\sqrt{-1}j\theta}.$$

Assume that both $A(\cdot)$ and $B(\cdot)$ are elements of $L_\infty(S^1, \mathbb{C}^{n\times n})$ and $L_\infty(S^1, \mathbb{C}^{n\times m})$, respectively. This guarantees that the associated multiplication operators

$$\begin{aligned}
\mathscr{A} &: L^2(S^1, \mathbb{C}^n) \longrightarrow L^2(S^1, \mathbb{C}^n), \quad x(e^{\sqrt{-1}\theta}) \mapsto A(e^{\sqrt{-1}\theta})x(e^{\sqrt{-1}\theta}),\\
\mathscr{B} &: L^2(S^1, \mathbb{C}^m) \longrightarrow L^2(S^1, \mathbb{C}^n), \quad u(e^{\sqrt{-1}\theta}) \mapsto B(e^{\sqrt{-1}\theta})u(e^{\sqrt{-1}\theta})
\end{aligned}$$

are bounded linear operators with operator norms

$$\begin{aligned}
\|\mathscr{A}\| &= \|A(\cdot)\|_\infty = \text{ess sup}_{0\le\theta\le 2\pi} \|A(e^{\sqrt{-1}\theta})\|,\\
\|\mathscr{B}\| &= \text{ess sup}_{0\le\theta\le 2\pi} \|B(e^{\sqrt{-1}\theta}).\|
\end{aligned}$$

It follows that the infinite-dimensional control system

$$\dot{x}(t) = \mathscr{A}x(t) + \mathscr{B}u(t) \tag{12.22}$$

that describes the platoon model (12.21) is well defined on the Hilbert space $L^2(S^1, \mathbb{C}^n)$. Note that the inputs for (12.22) assume values in the infinite-dimensional Hilbert space $L^2(S^1, \mathbb{C}^m)$. We present the following Hautus-type condition for approximate reachability.

Theorem 12.19. *The infinite platoon (12.21) is approximately reachable on the Hilbert space $L^2(S^1, \mathbb{C}^n)$ if and only if the reachability rank condition*

$$\operatorname{rk}\left(zI_n - A(e^{\sqrt{-1}\theta}),\, B(e^{\sqrt{-1}\theta})\right) = n$$

is satisfied for all $z \in \mathbb{C}$ and almost all $\theta \in [0, 2\pi]$.

Proof. By the isometric properties of the Fourier transform, both systems (12.21) and (12.22) are similar and thus have identical reachability properties. Characterizations of the approximate reachability of linear systems on a Hilbert space are well known; see, for example, the textbook by Curtain and Zwart (1995). In fact, approximate reachability in finite time $T > 0$ is guaranteed for (12.22) if and only if the image of the reachability operator $\mathscr{R}_T : L^2([0,T],X) \longrightarrow L^2(S^1,\mathbb{C}^n)$,

$$\mathscr{R}_T u = \int_0^T e^{(T-s)\mathscr{A}}\mathscr{B}u(s)ds,$$

is dense in $L^2(S^1,\mathbb{C}^n)$. Since the closure $\overline{\operatorname{Im}\mathscr{R}_T}$ of the image of $\mathscr{R}_T$ coincides with the kernel of the dual operator, we conclude that the approximate reachability of (12.21) is equivalent to the condition

$$\int_0^{2\pi} f(e^{\sqrt{-1}\theta})^* \exp(tA(e^{\sqrt{-1}\theta}))B(e^{\sqrt{-1}\theta})d\theta = 0 \quad \text{for all } t \geq 0 \implies f(e^{\sqrt{-1}\theta}) = 0.$$

Of course, this is equivalent to the familiar Kalman rank condition

$$\operatorname{rk}\left(B(e^{\sqrt{-1}\theta}),\ldots,A(e^{\sqrt{-1}\theta})^{nB}(e^{\sqrt{-1}\theta})\right) = n$$

for almost all $\theta \in [0,2\pi]$. Thus the result follows by applying the standard Hautus conditions for linear systems. ■

A crucial implication for the convergence dynamics of platoons, when working in the Hilbert space ℓ^2, is that all trajectories $(x_n(t))_n$ converge to zero as $n \to \pm\infty$. Of course, this is a very restrictive assumption that is often not desirable in practice. For such reasons, Feintuch and Francis (2012) started an investigation of platoons in the Banach space ℓ^∞ of bounded bi-infinite sequences in $\mathbb{R}$. We endow ℓ^∞ with the norm

$$\|x\|_\infty = \sup_{n\in\mathbb{Z}} |x_n|,$$

which turns ℓ^∞ into a Banach space. Since methods from Fourier analysis cannot be applied, the analysis becomes more difficult. We do not go into details here but refer the reader to Feintuch and Francis (2012) for a discussion of several simple cases.

3. Finite-Dimensional Control of Platoons. A further drawback of the preceding analysis is that it assumes an infinite-dimensional Hilbert space of input values. We now extend the reachability analysis of platoons to the more difficult, and more interesting, case of finite-dimensional controls.

We start with a simple example of a one-sided infinite platoon over $\mathbb{N}_0$ that was first described decades ago by Fuhrmann (1972). Consider the discrete-time control system in $\ell^2_+(\mathbb{C})$

$$x(t+1) = Sx(t) + bu(t), \tag{12.23}$$

where $b \in \ell_+^2(\mathbb{C})$ and

$$S : \ell_+^2(\mathbb{C}) \longrightarrow \ell_+^2(\mathbb{C}), \quad S(x_0, x_1, x_2, \ldots) := (0, x_0, x_1, \ldots),$$
$$S^* : \ell_+^2(\mathbb{C}) \longrightarrow \ell_+^2(\mathbb{C}), \quad S(x_0, x_1, x_2, \ldots) := (x_1, x_2, x_3, \ldots)$$

denotes the right shift operator and left shift operator, respectively. Thus S and S^* are bounded linear operators on $\ell_+^2(\mathbb{C})$. S is an isometry while S^* is a contraction. The study of these shift operators is due to Beurling (1949). The spectral properties of S and S^* are well known; for example, S does not have eigenvalues. The spectrum of both S and S^* is equal to the closed unit disc $\overline{\mathbb{D}}$. The set of eigenvalues of S^* is $\mathbb{D}$, and the associated eigenvector of each eigenvalue $\lambda \in \mathbb{D}$ is $(1, \lambda, \lambda^2, \lambda^3, \ldots)$. Equivalently, expressed in the Hardy space $H^2(\mathbb{D})$, the associated eigenfunction is

$$e_\lambda(z) = \frac{\sqrt{1 - |\lambda|^2}}{1 - \lambda z}.$$

Spelled out in the coordinates of the sequence $x(t)$, system (12.23) is equivalent to the 2D system $[x_{-1}(t) := 0]$

$$x_k(t+1) = x_{k-1}(t) + b_k u(t), \quad t, k \in \mathbb{N}_0.$$

Let

$$b(e^{\sqrt{-1}\theta}) = \sum_{k=0}^{\infty} b_k e^{k\sqrt{-1}\theta}$$

denote the discrete Fourier transform of b. Since $b \in \ell_+^2(\mathbb{C})$, the Fourier transform $b(\cdot)$ extends to a holomorphic function in $\mathbb{D}$ and $b(z) \in H^2(\mathbb{D})$. Note that if $b(z)$ is a rational function with no poles and zeros in the closed unit disc, then b is an outer function. The following result by Beurling (1949) appears as Lemma 4.1 in Fuhrmann (1972). Since $b(z) = \frac{2}{2-z} = \sum_{k=0}^{\infty} 2^{-k} z^k$ is outer, the result implies that the platoon system $[x_{-1}(t) := 0]$

$$x_k(t+1) = x_{k-1}(t) + 2^{-k} u(t), \quad t, k \in \mathbb{N}_0.$$

is approximately reachable.

Theorem 12.20. *System (12.23) is approximately reachable in $\ell_+^2(\mathbb{C})$ if and only if b is an outer function.*

Proof. For the convenience of the reader we recall the main arguments from Fuhrmann (1972). System (12.23) is approximately reachable at zero if and only if the functions $e^{\sqrt{-1}k\theta} b(e^{\sqrt{-1}k\theta})$ span H^2. The span V is invariant under multiplication by $e^{\sqrt{-1}k\theta}$, and therefore Beurling's theorem implies that $V = qH^2$

for an inner function q. Thus $b \in qH^2$, i.e., q divides b. Thus $V = H^2$ if and only if b does not contain a nontrivial inner function. By the inner-outer factorization theorem, this is equivalent to b being outer. ∎

The situation becomes quite different if we replace the forward shift operator S in (12.23) with the backward shift S^*. The approximate reachability of

$$x(t+1) = S^*x(t) + bu(t) \tag{12.24}$$

is equivalent to b being a cyclic vector for S^*. The question of characterizing cyclic vectors for the backward shift was first raised and answered by Douglas, Shapiro and Shields (1970) and extended to the multivariable case in Fuhrmann (1976b). Their characterization is, however, not as simple as that for S. A sufficient condition by Douglas, Shapiro and Shields (1970) for the cyclicity of b for the backward shift is that the Fourier transform $b(z) \in H^2$ can be analytically continued across all points of an arc in S^1, with the exception of an isolated branch point in the arc. An example of a cyclic vector is $b(z) = \exp(1/(z-2))$. No rational function $b(z) \in H^2$ is cyclic, and therefore rational stable functions lead to a nonreachable system (12.24). In fact, from Kronecker's theorem 4.18 one can deduce that the span of the orbit $(S^*)^{n}b$ is finite-dimensional.

It is possible to extend the analysis to broader classes of one-sided platoon models. Let $a(z) = \sum_{j=0}^{\infty} a_j z^j$ denote an analytic function in $H^2(\mathbb{D})$, with $\sup_{|z|\leq 1} |a(z)| < \infty$, and let $b(z) = \sum_{j=0}^{\infty} b_j z^j \in H^2(\mathbb{D})$. Then the discrete-time single-input linear control system on $\ell_+^2(\mathbb{C})$ is well defined as

$$x(t+1) = a(S)x(t) + bu(t) \tag{12.25}$$

or, equivalently, as

$$x_k(t+1) = \sum_{j=0}^{k} a_{k-j} x_j(t) + b_k u(t), \quad t,k \in \mathbb{N}_0.$$

Note that for $a(z) = z$ this specializes to (12.23). For the proof of the subsequent theorem, we apply methods from the theory of composition operators in Hardy spaces; see, for example, Douglas, Shapiro and Shields (1970) and Shapiro (1993) for further details. In particular, we make use of the following classical result by Walsh (1965) for polynomial approximations.

Theorem 12.21 (Walsh). *Let $a : \mathbb{D} \longrightarrow \mathbb{C}$ be an injective holomorphic function such that the boundary of $a(\mathbb{D})$ is a Jordan curve. Then the set $\{p \circ a \mid p \in \mathbb{C}[z]\}$ of polynomials in a is dense in $H^2(\mathbb{D})$.*

The discussion of the following two examples is taken from Bourdon and Shapiro (1990).

Example 12.22. The univalent function $a(z) = \frac{z}{2-z} \in H^\infty(\mathbb{D})$ maps the closed disc $\overline{\mathbb{D}}$ conformally into itself with fixed points $0, 1$. Note that $\sup_{|z|\leq 1} |a(z)| = 1$. The composition operator $C_a : H^2(\mathbb{D}) \longrightarrow H^2(\mathbb{D})$ is bounded but does not have cyclic vectors. Therefore, system (12.25)

$$x_k(t+1) = \sum_{j=0}^{k-1} 2^{j-k} x_j(t) + b_k u(t), \quad t, k \in \mathbb{N}_0,$$

is for no $b \in \ell^2_+(\mathbb{C})$ approximately reachable in $\ell^2_+(\mathbb{C})$. On the other hand, the univalent function $a(z) = \frac{1}{2-z} \in H^\infty(\mathbb{D})$ has $z = 1$ as its only fixed point and satisfies $\sup_{|z|\leq 1} |a(z)| = 1$. The composition operator $C_a : H^2(\mathbb{D}) \longrightarrow H^2(\mathbb{D})$ is bounded and cyclic. Thus there exists a generic set of elements $b \in \ell^2_+(\mathbb{C})$ such that

$$x_k(t+1) = \sum_{j=0}^{k} 2^{j-k-1} x_j(t) + b_k u(t), \quad t, k \in \mathbb{N}_0,$$

is approximately reachable in $\ell^2_+(\mathbb{C})$.

The following generalization of Theorem 12.20 is a simple consequence of the Walsh theorem.

Theorem 12.23. *Let $b \in \ell^2_+(\mathbb{C})$. Assume that $a \in H^2(\mathbb{D})$ defines an injective analytic function $a : \mathbb{D} \longrightarrow \mathbb{C}$ such that the boundary of $a(\mathbb{D})$ is a Jordan curve. Assume further that $\sup_{|z|\leq 1} |a(z)| < 1$. System (12.25) is approximately reachable in $\ell^2_+(\mathbb{C})$ if and only if b is an outer function.*

Proof. The reachable set V consists of all functions $(p \circ a)b$, where $p \in \mathbb{C}[z]$ is a polynomial. By Theorem 12.21 of Walsh, this implies that the closure $\overline{V}$ of the reachable set in H^2 is equal to bH^2. Let $b = fg$, with f inner and g outer, denote the inner-outer factorization of b. Then the closure of the reachable set $bH^2 = fH^2$ is equal to H^2 if and only if f is constant, i.e., if and only if b is outer. ∎

Actually, a stronger version of Walsh's theorem is valid asserting that the polynomials in a are uniformly dense in the space of uniformly continuous bounded analytic functions on $\mathbb{D}$. This implies approximate reachability in $\ell^2_+(\mathbb{C})$ with respect to the sup-norm on H^2. We next proceed to show how one can apply such uniform approximation techniques in a more general context. Consider bi-infinite platoons with spatial domain $\mathbb{Z}$, where m controls $u(t) = (u_1(t), \ldots, u_m(t))^\top \in \mathbb{R}^m$ are broadcasted to the platoon using

$$u_k(t) = \beta_k u(t),$$

where $\beta_k \in \mathbb{R}^{m\times m}$. This leads to the **broadcast control platoon system**

$$\dot{x}_k(t) = \sum_{j=-\infty}^{\infty} A_j x_{k-j}(t) + \sum_{j=-\infty}^{\infty} B_j \beta_{k-j} u(t), \quad k \in \mathbb{Z}.$$

We assume that the bi-infinite sequence of matrices $\beta = (\beta_k)$ defines an L^2 Fourier transform

$$\beta(e^{\sqrt{-1}\theta}) := \sum_{k=-\infty}^{\infty} \beta_{ke}^{-\sqrt{-1}k\theta} \in L^2(S^1;\mathbb{C}^{m\times m}).$$

The assumptions on A and B are as previously. Define matrix-valued functions as

$$A(e^{\sqrt{-1}\theta}) = \sum_{k=-\infty}^{\infty} A_{ke}^{-\sqrt{-1}k\theta}, \quad \hat{B}(e^{\sqrt{-1}\theta}) = B(e^{\sqrt{-1}\theta})\beta(e^{\sqrt{-1}\theta}).$$

The associated multiplication operators

$$\mathscr{A} : L^2(S^1,\mathbb{C}^n) \longrightarrow L^2(S^1,\mathbb{C}^n) \quad \text{and} \quad \hat{\mathscr{B}} : \mathbb{C}^m \longrightarrow L^2(S^1,\mathbb{C}^n)$$

are bounded linear operators, whereas the input-state operator $\hat{\mathscr{B}}$ is finite-dimensional and therefore defines a compact operator. We thus obtain the system on the Hilbert space $L^2(S^1;\mathbb{C}^n)$ with finite-dimensional controls as

$$\dot{x}(t) = \mathscr{A}x(t) + \hat{\mathscr{B}}u(t). \tag{12.26}$$

The associated finite-dimensional, parameter-dependent, linear system on $\mathbb{C}^n$ is

$$\dot{z}(t) = \mathscr{A}(e^{\sqrt{-1}\theta})z(t) + \mathscr{B}(e^{\sqrt{-1}\theta})\beta(e^{\sqrt{-1}\theta})u(t),$$

with parameter space S^1 being the unit circle. For a compact subset $\mathbf{P} \subset S^1$ consider the Hilbert spaces

$$L^2_{\mathbf{P}}(\mathbb{C}^n) = \{f \in L^2(S^1;\mathbb{C}^n) \mid f = 0 \text{ outside } \mathbf{P}\}, \quad \ell^2_{\mathbf{P}} = \mathscr{F}^{-1}(L^2_{\mathbf{P}}),$$

with the isomorphism of Hilbert spaces $L^2_{\mathbf{P}}(\mathbb{C}^n) \simeq L^2(\mathbf{P},\mathbb{C}^n)$.

By extending square-integrable matrix-valued functions $A(e^{\sqrt{-1}\theta})$, $B(e^{\sqrt{-1}\theta})$, and $\beta(e^{\sqrt{-1}\theta})$ on $\mathbf{P}$ by zero to functions on S^1 one can identify these matrix functions with unique elements of, for example, $L^2_{\mathbf{P}}(\mathbb{C}^{n\times n})$. Then the linear multiplication operators $\mathscr{A}$ and $\hat{\mathscr{B}}$ map $L^2_{\mathbf{P}}(\mathbb{C}^n)$ and $\mathbb{R}^m$ into $L^2_{\mathbf{P}}(\mathbb{C}^n)$. In particular, the linear system (12.26) is restricted to a control system on the Hilbert space $L^2_{\mathbf{P}}(\mathbb{C}^n)$. In the single-input case we obtain the following approximate reachability result.

Theorem 12.24. *Let $m = 1$, and let $\boldsymbol{P} \neq [0, 2\pi]$ denote a nonempty compact interval contained in $[0, 2\pi]$. Assume that the matrix-valued functions $\theta \mapsto A(e^{\sqrt{-1}\theta})$, $B(e^{\sqrt{-1}\theta})$, and $\beta(e^{\sqrt{-1}\theta})$ are continuous on $\boldsymbol{P}$. Assume further that the following conditions are satisfied:*

1. *The pair $(A(e^{\sqrt{-1}\theta}), B(e^{\sqrt{-1}\theta})\beta(e^{\sqrt{-1}\theta}))$ is reachable for all $\theta \in \mathbb{P}$.*
2. *The spectra of $A(e^{\sqrt{-1}\theta})$ and $A(e^{\sqrt{-1}\theta'})$ are disjoint for each $\theta, \theta' \in \boldsymbol{P}, \theta \neq \theta'$.*
3. *The eigenvalues of $A(e^{\sqrt{-1}\theta})$ are simple for each $\theta \in \boldsymbol{P}$.*

Then the restricted system (12.26) on the Hilbert subspace $L^2_{\mathbf{P}}(\mathbb{C}^n)$ is approximately reachable in finite time $T > 0$.

Proof. Let $b(e^{\sqrt{-1}\theta}) := B(e^{\sqrt{-1}\theta})\beta(e^{\sqrt{-1}\theta})$. Theorem 12.6 implies that the parameter-dependent system

$$\frac{\partial x(t,\theta)}{\partial t} = A(e^{\sqrt{-1}\theta})x(t,\theta) + b(e^{\sqrt{-1}\theta})u(t), \quad \theta \in \mathbf{P},$$

is uniformly ensemble reachable in finite time. Note that

$$\int_0^T \exp((T-s)\mathscr{A})\hat{\mathscr{B}}u(s)ds$$

coincides with the $L^2_{\mathbf{P}}$ function

$$\theta \mapsto \int_0^T \exp\left((T-s)A(e^{\sqrt{-1}\theta})\right) b(e^{\sqrt{-1}\theta})u(s)ds.$$

This implies that the image of the reachability operator $\mathscr{R}_T : L^2([0,T];\mathbb{C}^m) \longrightarrow L^2(\mathbf{P}, \mathbb{C}^n) = L^2_{\mathbf{P}}(\mathbb{C}^n)$,

$$\mathscr{R}_T(u) = \int_0^T \exp((T-s)\mathscr{A})\hat{\mathscr{B}}u(s)ds,$$

is dense in $L^2_{\mathbf{P}}(\mathbb{C}^n)$. This completes the proof. ■

12.4 Control of Partial Differential Equations

In this section we explore several instances where the control of PDEs interacts with networks of systems. This includes the realization of interconnected systems and platoons as discretizations of PDEs and showing how results from parametric linear systems can be used to gain further insight into PDEs using Fourier-transform techniques. Finally, we explain how one can control the state-space probability distributions of linear systems by solving the associated control task for the Liouville equation.

1. Networks as Discretizations of the Heat Equation. Let us consider one of the simplest PDEs, the classical **heat equation** in one spatial variable $z \in [0,1]$, with boundary controls

$$\frac{\partial \psi(t,z)}{\partial t} = \frac{\partial^2 \psi(t,z)}{\partial z^2},$$
$$\psi(0,x) = 0, \quad \psi(t,0) = u_0(t), \quad \psi(t,1) = u_1(t).$$

Here the boundary value functions $u_0(t)$ and $u_1(t)$ are regarded as control variables. For each nonnegative integer N and step size $h = \frac{1}{N}$ we subdivide the domain $[0,1]$ into $N+1$ equidistant points $z_i = ih, i = 0,\ldots,N$. Consider a lumped discretization as $x(t) = (\psi(t,\frac{1}{N}),\ldots,\psi(t,\frac{N-1}{N}))^\top \in \mathbb{R}^{N-1}$, and assume that the boundary value functions $u(t) = (u_0(t),u_1(t))^\top$ are known. Then the boundary value condition $\psi(0,x) = 0$ corresponds to the initial condition $x(0) = 0$. Using standard Taylor approximations we can replace the second-order differential operator $\frac{\partial^2 \psi(t,z)}{\partial z^2}$ with its associated difference operator as

$$\frac{\psi(t,z+h) - 2\psi(t,z) + \psi(t,z-h)}{h^2}.$$

By neglecting second-order error terms we end up with the discretized form of the heat equation as

$$\dot{x}(t) = \frac{1}{h^2}\left(Ax(t) + Bu(t)\right). \tag{12.27}$$

Of course, after rescaling time in x,u via $x(h^2t), u(h^2t)$ this system becomes equivalent to the linear system $\dot{x} = Ax + Bu$. Here $(A,B) \in \mathbb{R}^{(N-1)\times(N-1)} \times \mathbb{R}^{(N-1)\times 2}$ are the reachable pair

$$A = \begin{pmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & \ddots & \ddots & & \\ & & \ddots & \ddots & 1 & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix}. \tag{12.28}$$

Note that (12.27) is an interconnected system for $N-1$ identical first-order integrators $\dot{x}_i = v_i$, coupled by interconnection matrices (12.28) for the states and inputs, respectively. Note further that A defines a Laplacian matrix for a tree, and its spectrum was analyzed in Chapter 7. It is easily seen that the discretized heat equation (12.27) is reachable.

As another example, consider the control of the heat equation on the unit circle, formulated as a periodic boundary value problem as

$$\begin{aligned} \frac{\partial \psi(t,z)}{\partial t} &= \frac{\partial^2 \psi(t,z)}{\partial z^2} + u(t)g(z), \\ \psi(t,0) &= \psi(t,1), \end{aligned}$$

where $g(z) = g(z+1)$ is assumed to be periodic with period one. By discretizing this system using $x_N = x_0, x_{N+1} = x_1$ and setting $g(ih) = b_i$ we obtain the single-input control system $\dot{x} = \frac{1}{h^2}(Ax + bu)$ with

$$A = \begin{pmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & \ddots & \ddots & & \\ & & \ddots & \ddots & 1 & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_N \end{pmatrix}. \tag{12.29}$$

Here, A is a circulant matrix, and the reachability of such systems was analyzed in Chapter 9 using the module structure on the set of circulant matrices. We refer to Brockett and Willems (1974) for further discussion of discretized PDEs from a module-theoretic point of view. Note that (12.29) is, up to sign, identical with the finite platoon system (12.19). In particular, Theorem 12.18 implies that (12.29) is not reachable.

The discretization process can be applied to more general boundary value problems. We illustrate this by the following variant of the controlled heat equation

$$\begin{aligned} &\frac{\partial \psi(t,z)}{\partial t} = \frac{\partial^2 \psi(t,z)}{\partial z^2}, \quad (t,z) \in [0,T] \times [0,1], \\ &\frac{\partial \psi(t,0)}{\partial z} = 0, \quad \frac{\partial \psi(t,1)}{\partial z} = u(t). \end{aligned}$$

We refer the reader to Chapter 2.5.3 in the book by Coron (2007) for a detailed discussion of the reachability properties of this system, including an approximate reachability result Theorem 2.76 that is derived using flatness techniques. Using Taylor approximations (with step sizes identical to those for the second derivative term),

$$\frac{\psi(t,h) - \psi(t,0)}{h} \quad \text{and} \quad \frac{\psi(t,1) - \psi(t,1-h)}{h},$$

for $\frac{\partial \psi(t,0)}{\partial z}$ and $\frac{\partial \psi(t,1)}{\partial z}$, respectively, the boundary conditions become $\psi(t,h) = \psi(t,0)$ and $\psi(t,1) = \psi(t,1-h) + hu(t)$. Thus we arrive at the reachable single-input system

$$\dot{x}(t) = \frac{1}{h^2}\left(Ax(t) + hbu(t)\right),$$

with

$$A = \begin{pmatrix} -1 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & \ddots & \ddots & & \\ & & \ddots & \ddots & 1 & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 1 \end{pmatrix}.$$

Again, this system is reachable and can be rescaled to standard form $\dot{x} = Ax + bu$. The spectral properties of matrix A were discussed in Section 8.7.

Instead of assuming a compact spatial domain, one can also consider the heat equation on an unbounded domain. This leads to infinite platoons of systems as their discretizations. Thus, consider, for example, the controlled heat equation on the nonnegative real line $[0,\infty)$:

$$\begin{aligned} \frac{\partial \psi(t,z)}{\partial t} &= \frac{\partial^2 \psi(t,z)}{\partial z^2} + g(z)u(t), \\ \psi(0,z) &= 0. \end{aligned}$$

Consider a lumped approximation $x(t) = (x_k(t)) = (\psi(t,k))_{k\in\mathbb{N}_0} \in \ell^2_+$ of ψ, and $(b_k)_{k\in\mathbb{N}_0} := (g(k))_{k\in\mathbb{N}_0} \in \ell^2_+$ with step size $h = 1$. Using the standard discretization for the second-order derivative one obtains the infinite platoon

$$\dot{x}_k(t) = x_{k+1}(t) - 2x_k(t) + x_{k-1}(t)(t) + b_{ku}(t), \quad k \in \mathbb{N}_0.$$

This shows that the study of infinite platoons can be of use for the control of PDEs.

The preceding examples illustrate that interesting interconnection matrices arise as discretizations of PDEs. One can take this idea a step further by considering discretizations of parametric linear systems coupled by diffusive terms. This leads to networks of linear systems. Thus, consider, for example, the family of linear systems

$$\begin{aligned} \frac{\partial \psi(t,z)}{\partial t} &= \alpha\psi(t,z) + \beta v(t,z), \\ y(t,z) &= \gamma\psi(t,z), \\ v(t,z) &= \frac{\partial^2 y(t,z)}{\partial^2 z} + Bu(t), \end{aligned}$$

where the SISO system $(\alpha,\beta,\gamma)\in\mathbb{R}^{n\times n}\times\mathbb{R}^n\times\mathbb{R}^{1\times n}$ is reachable and observable, $z\in[0,1]$ and $\psi(t,z)\in\mathbb{R}^n$. By discretizing the second-order differentiation term as before we arrive at the interconnected linear system

$$\dot{x}=(I_N\otimes\alpha+A\otimes\beta\gamma)x+(B\otimes\beta)u(t),$$

where A is defined in (12.28) and B is arbitrary. By Theorem 9.15, this network is reachable whenever (A,B) is reachable. If the spatial domain $[0,1]$ is replaced by the real line $\mathbb{R}$, this leads to infinite networks of linear systems.

2. Fourier-Transform Techniques. The preceding examples showed how large-scale interconnected control systems arise naturally as discretizations of boundary control problems for PDEs. Using Fourier-transform techniques, one can associate families of control systems to such PDEs on $\mathbb{R}^d$. To illustrate this idea, we consider the controlled heat equation in one spatial variable on the real line $\mathbb{R}$:

$$\begin{aligned}\frac{\partial\psi(t,z)}{\partial t}&=a\psi(t,z)+\frac{\partial^2\psi(t,z)}{\partial z^2}+u(t)\sqrt{\frac{2}{\pi}}\frac{\sin(Rz)}{z},\\ \psi(0,z)&=0.\end{aligned}\tag{12.30}$$

Our goal is to find a control function $u(t)$ that steers the initial temperature distribution $\psi(0,z)=0$ to a final distribution $\psi(T,z)=\psi^*(z)$ in finite time $T>0$. More specifically, for $\varepsilon>0$ and a function ψ^* in the Sobolev space $W^2(\mathbb{R})$ we want to find $T>0$ and a control $u:[0,T]\longrightarrow\mathbb{R}$ such that

$$\int_{-\infty}^{\infty}(\psi(T,z)-\psi^*(z))^2\,dz<\varepsilon.\tag{12.31}$$

We refer to this as the **approximate reachability task** for ψ^* in (12.30). We emphasize that the control function in our problem is independent of the spatial variable z. Thus the freedom one has in controlling the system is quite limited. To approach such reachability questions, it is useful to note that (12.30) has a unique solution $\psi(t,z)\in C^1([0,\infty),W^2(\mathbb{R})\cap W_0^1(\mathbb{R}))$ for piecewise smooth input functions $u(t)$; see Pazy (1983), Corollary 7.2.8. Thus the control problem is meaningful only by assuming $\psi^*\in W^2(\mathbb{R})\cap W_0^1(\mathbb{R})$. Since the space $C_0^\infty(\mathbb{R})$ of smooth functions with compact support is dense in $W^2(\mathbb{R})$, one has $W_0^1(\mathbb{R})=W^2(\mathbb{R})$. Thus the control problem for the heat equation (12.30) on the real axis takes place in $W^2(\mathbb{R})$. Our approach now is to replace the heat equation by an ordinary differential equation using the Fourier transform. This enables us to show that, under suitable assumptions on ψ^*, such approximate reachability tasks for PDEs are related to the ensemble control problem for parameter-dependent linear systems.

Recall that the **Fourier transform** on $\mathbb{R}$ defines the linear isometry $\mathscr{F}: L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$,

$$(\mathscr{F}\psi)(\theta) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\sqrt{-1}\theta z} \psi(z) dz,$$

with inverse transform

$$(\mathscr{F}^{-1} f)(z) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\sqrt{-1}\theta z} f(\theta) d\theta.$$

It is well known that $\mathscr{F}$ maps the Sobolev space $W^2(\mathbb{R})$ exactly onto the space of all functions $\phi(\theta)$ such that $(1+|\theta|)^2\phi(\theta) \in L^2(\mathbb{R})$. Let

$$H(x) = \begin{cases} 1 & \text{for} \quad x \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

denote the Heaviside function. Note that for each $R > 0$ the Fourier transform of $\sqrt{\frac{2}{\pi}} \frac{\sin(Rz)}{z}$ is equal to $H(R - |z|)$. Thus, by Fourier-transforming equation (12.30), we obtain for $x(t, \theta) := (\mathscr{F}\psi)(t, \theta)$ the parameter-dependent control system

$$\dot{x} = (a - \theta^2)x(t, \theta) + H(R - |\theta|)u(t), \quad x(0, \theta) = 0. \tag{12.32}$$

By restricting ourselves to the compact parameter domain $\mathbf{P} = [0, R]$ we obtain the linear parameter-dependent system

$$\dot{x} = (a - \theta^2)x(t, \theta) + u(t), \quad x(0, \theta) = 0. \tag{12.33}$$

Since the Fourier transform defines an isometry on $L^2(\mathbb{R})$, the approximate reachability of the PDE (12.30) on the spatial domain $\mathbb{R}$ is equivalent to the existence of an input function $u(t)$ that is *independent of the parameter* $\theta \in [0, R]$ and has the following approximation property: For each function $\theta \mapsto x^*(\theta)$ of terminal states in $\mathscr{F}(W^2)$ and $\varepsilon > 0$ there exists $T > 0$ such that the solution of (12.30) satisfies

$$\int_0^R (x(T, \theta) - x^*(\theta))^2 d\theta < \varepsilon.$$

Thus we see that the approximate reachability task for the heat equation is equivalent to L^2-ensemble reachability for a simple family of first-order linear systems. Our previous results on uniform ensemble reachability apply and yield corresponding results on approximate reachability for PDEs. However, there is a "but" insofar as our results will only imply reachability within a certain frequency band. These observations are in harmony with stronger positive approximate reachability results established for the heat equation (12.30) on arbitrary bounded domains; see Theorem 2.76 in the book by Coron (2007).

Theorem 12.25. *Assume that $\psi^*(z) \in W^2$ is such that its Fourier transform $x^*(\theta) = \mathscr{F}(\psi^*)(\theta)$ is an even function of θ with support contained in $[-R, R]$. Then the heat equation (12.30) is approximately reachable to ψ^*.*

Proof. We first show that the approximate reachability of the heat equation to ψ^* is equivalent to the L^2-ensemble reachability of the linear parametric system (12.33) to x^*. By Theorem 12.6, the family (12.33) is uniform ensemble reachable on the parameter interval $[0, R]$. This then completes the proof. To prove the equivalence of the two notions of reachability, suppose that u is an input such that $\psi(T, \cdot) \in W^2$ satisfies (12.31). Then $x(t, \theta) = \mathscr{F}(\psi(t, \cdot)$ is a solution of (12.32). Since $\mathscr{F}$ is an isometry on L^2, we obtain for $x^* := \mathscr{F}(\psi^*)$

$$\int_{-\infty}^{\infty} (\psi(T,z) - \psi^*(z))^2 \, dz = \int_{-\infty}^{\infty} (x(T,\theta) - x^*(\theta))^2 \, d\theta = \int_{-R}^{R} (x(T,\theta) - x^*(\theta))^2 \, d\theta$$
$$= 2\int_{0}^{R} (x(T,\theta) - x^*(\theta))^2 \, d\theta < 2\varepsilon.$$

Here the second equation follows from the support property of x^*, while the third follows from the evenness of the functions $x(T, \theta)$ and $x^*(\theta)$. Conversely, for each solution $x(t, \theta)$ of (12.33) we note that $H(R - |\theta|)x(t, \theta)$ defines a solution of (12.32). Since $\theta \mapsto H(R - |\theta|)x(t, \theta)$ has compact support, $(1 + |\theta|)^2 H(R - |\theta|)x(t, \theta)$ also has compact support. Thus $H(R - |\theta|)x(t, \theta)$ is an element of $\mathscr{F}(W^2(\mathbb{R}))$ for all $t \geq 0$. This shows that the inverse Fourier transform $\mathscr{F}^{-1}(H(R - |\theta|)x(t, \theta))$ maps solutions to (12.32) bijectively to the solutions of the heat equation (12.30). ∎

3. Control of Liouville Equation. We now present a reachability result by Brockett on the Liouville equation, restricted to the space of Gaussian distributions. We begin by deriving the classical Liouville transport equation for time-varying vector fields. Let $f(x,t)$ be a time-varying complete C^k-vector field on a smooth orientable Riemannian manifold M. Let dx denote the canonical volume form on M and $\rho(0, \cdot)$ denote a smooth function with

$$\int_M \rho(0,x)dx = 1.$$

Let $\phi_{t,s}$ denote the flow semigroup of f. Thus $\phi_{t,s}$ is a diffeomorphism on M with $\phi_{t,s}\phi_{s,t} = id_M$ and

$$\frac{\partial \phi_{t,s}(x)}{\partial t} = f(\phi_{t,s}(x), t).$$

Define

$$\rho(t,x) = \rho(0, \phi_{0,t}(x)) \det D\phi_{0,t}(x),$$

i.e., $\rho(t,x)dx$ is the pullback of the volume form $\rho(0,x)dx$ by the diffeomorphism $\phi_{0,t}$. Applying the transformation theorem one concludes that, for all t,

$$\int_M \rho(t,x)dx = 1.$$

Consider a smooth function $\psi(x)$ on M with compact support. By a change of variables,

$$\int_M \psi(x)\rho(t,x)dx = \int_M \psi(x)\rho(0,\phi_{0,t}(x))\det D\phi_{0,t}(x)dx = \int_M \psi(\phi_{t,0}(x))\rho(0,x)dx,$$

and so, by differentiating both sides, we obtain

$$\begin{aligned}\int_M \psi(x)\frac{\partial\rho(t,x)}{\partial t}dx &= \frac{d}{dt}\int_M \psi(x)\rho(t,x)dx = \int_M \frac{\partial\psi(\phi_{t,0}(x))}{\partial t}\rho(0,x)dx\\ &= \int_M d\psi(\phi_{t,0}(x))f(\phi_{t,0}(x),t)\rho(0,x)dx\\ &= \int_M d\psi(x)f(x,t)\rho(0,\phi_{0,t}(x))\det D\phi_{0,t}(x)dx\\ &= \int_M d\psi(x)f(x,t)\rho(t,x)dx.\end{aligned}$$

For time-varying vector fields F, the divergence on M satisfies the well-known identity $\operatorname{div}(\psi(x)F(x,t)) = d\psi(x)F(x,t) + \psi(x)\operatorname{div}F(x,t)$. Applying the divergence theorem, we obtain, for $F(x,t) = \rho(t,x)f(x,t)$,

$$\begin{aligned}\int_M \psi(x)\frac{\partial\rho(t,x)}{\partial t}dx &= -\int_M \psi(x)\operatorname{div}(f(x,t)\rho(t,x))dx + \int_M \operatorname{div}(\psi(x)f(x,t)\rho(t,x))dx\\ &= -\int_M \psi(x)\operatorname{div}(f(x,t)\rho(t,x))dx.\end{aligned}$$

Thus we conclude that $\rho(t,x)$ satisfies the transport equation

$$\frac{\partial\rho(t,x)}{\partial t} = -\operatorname{div}(f(x,t)\rho(t,x)), \quad \rho(0,x) = \rho(x). \tag{12.34}$$

Conversely, if ρ is a C^k-function on $\mathbb{R}^n$, then $\rho(t,x) = \rho(\phi_{0,t}(x))\det D\phi_{0,t}(x)$ is the unique C^k-solution of the initial value problem (12.34). Applying this equation to a control affine vector field $f(x)+ug(x)$, a major distinction between open-loop and closed-loop control becomes manifest. In fact, for open-loop controls we obtain a control system on the space of density functions as

$$\frac{\partial\rho(t,x)}{\partial t} = -\operatorname{div}(f(x)\rho(t,x)) - u(t)\operatorname{div}(g(x)\rho(t,x)),$$

while smooth feedback control leads to

$$\frac{\partial \rho(t,x)}{\partial t} = -\operatorname{div}(f(x)\rho(t,x)) - u(x)\operatorname{div}(g(x)\rho(t,x)) - du(x)g(x)\rho(t,x). \tag{12.35}$$

The difference appears already for linear systems $f(x,u) = Ax + Bu$ in $\mathbb{R}^n$. We obtain the open-loop transport equation

$$\frac{\partial \rho(t,x)}{\partial t} = -\operatorname{tr}(A)\rho(t,x) - (Ax+Bu)^\top \nabla\rho(t,x), \tag{12.36}$$

while the closed-loop state feedback $u = Kx$ leads to

$$\frac{\partial \rho(t,x)}{\partial t} = -\operatorname{tr}(A+BK)\rho(t,x) - ((A+BK)x)^\top \nabla\rho(t,x). \tag{12.37}$$

The open-loop solution of (12.36) is

$$\rho(t,x) = e^{-t\operatorname{tr}A}\rho\left(0, e^{-tA}(x - \int_0^t e^{(t-s)A}Bu(s)ds)\right),$$

while the closed-loop solution of (12.37) under the state feedback $u = Kx$ is

$$\rho(t,x) = e^{-t\operatorname{tr}(A+BK)}\rho(0, e^{-t(A+BK)}x).$$

A **Gaussian distribution** function with positive definite covariance matrix $Q = Q^\top > 0$ and mean value $\mu \in \mathbb{R}^n$ is defined as

$$g_{Q,\mu}(x) = \frac{1}{\sqrt{(2\pi)^n \det Q}} \exp\left(-\frac{1}{2}(x-\mu)^\top Q^{-1}(x-\mu)\right).$$

Let $\mathscr{P}$ denote the convex set of positive definite real symmetric $n \times n$ matrices Q. Since the map $(Q,\mu) \mapsto g_{Q,\mu}$ is injective, we see that the set of Gaussian distributions forms a smooth manifold $\mathscr{G}$ that is diffeomorphic to $\mathscr{P} \times \mathbb{R}^n$ and is embedded into $C^\infty(\mathbb{R}^n)$. Moreover, the Gaussian distribution satisfies, for every invertible transformation $S \in GL_+(n,\mathbb{R})$ with positive determinant and $\mu \in \mathbb{R}^n$, the identity

$$g_{SQS^\top, S\mu}(x) = \frac{1}{\det S} g_{Q,\mu}(S^{-1}x).$$

We now consider the effect of transporting a Gaussian distribution by a linear control system. The solutions of the linear system $\dot{x} = Ax + Bu$ are

$$x(t) = \phi_{t,0}(x) = e^{tA}x + \int_0^t e^{(t-s)A}Bu(s)ds,$$

with

$$\phi_{0,t}(x) = e^{-tA}x - \int_0^t e^{-sA}Bu(s)ds.$$

Therefore,

$$g_{Q,\mu}(\phi_{0,t}(x)) = e^{t\,\mathrm{tr}(A)}g_{Q(t),\mu(t)}(x),$$

with

$$Q(t) = e^{tA}Qe^{tA^\top}, \quad \mu(t) = \phi_{t,0}(\mu) = e^{tA}\mu + \int_0^t e^{(t-s)A}Bu(s)ds.$$

This implies the explicit formula for the L^2-distance

$$\int_{\mathbb{R}^n}(g_{Q,\mu}(x) - g_{Q(t),\mu(t)}(x))^2 dx = \frac{1}{\sqrt{(4\pi)^n \det Q}} + \frac{e^{-t\,\mathrm{tr}A}}{\sqrt{(4\pi)^n \det Q}}$$
$$- \frac{2\exp\left(-\frac{1}{2}(\mu - \mu(t))^\top (Q + e^{tA}Qe^{tA^\top})^{-1}(\mu - \mu(t))\right)}{\sqrt{(2\pi)^n \det(Q + e^{tA}Qe^{tA^\top})}}.$$

Now consider the bilinear control system

$$\dot{x} = (A + BK(t))x + Bu(t),$$

where $K(t)$ and $u(t)$ act as independent control functions. Thus, by this process, we combine open-loop and closed-loop controls for the transport equation (12.39) on Gaussian density functions. This induces the control system on the parameter space for Gaussian densities as

$$\begin{aligned}\dot{Q}(t) &= (A + BK(t))Q(t) + Q(t)(A + BK(t))^\top,\\ \dot{\mu}(t) &= (A + BK(t))\mu(t) + Bu(t),\end{aligned} \tag{12.38}$$

which is equivalent to the restriction of (12.39) on $\mathscr{G}$.

Theorem 12.26 (Brockett (2012)). *Assume that (A,B) is reachable. Then the control system (12.38) on $\mathscr{P} \times \mathbb{R}^n$ is reachable. For pairs (Q_0, μ_0) and (Q_1, μ_1) in $\mathscr{P} \times \mathbb{R}^n$ there exists $T > 0$ and control functions $u(t)$ and $K(t)$ on $[0,T]$ that steer (12.38) from (Q_0, μ_0) to (Q_1, μ_1).*

Proof. The tangent space of $\mathscr{P}$ at Q consists of all matrices of the form $LQ + QL^\top$, with $L \in \mathbb{R}^{n\times n}$. Thus the solutions $Q(t)$ and $\mu(t)$ of (12.38) exist for all time and stay in $\mathscr{P} \times \mathbb{R}^n$. Every positive definite matrix Q can be factored as $Q = XX^\top$, with a factor $X \in GL_+(n,\mathbb{R})$ of positive determinant. Thus it suffices to prove reachability for the lifted system on $GL_+(n,\mathbb{R}) \times \mathbb{R}^n$:

$$\begin{aligned} \dot{X}(t) &= (A + BK(t))X(t), \\ \dot{\mu}(t) &= (A + BK(t))\mu(t) + Bu(t). \end{aligned}$$

This in turn is reachable if and only if the decoupled system

$$\begin{aligned} \dot{X}(t) &= (A + BK(t))X(t), \\ \dot{\mu}(t) &= A\mu(t) + Bu(t) \end{aligned}$$

is reachable. The second subsystem is linear and, thus, by assumption on (A,B), is reachable on $\mathbb{R}^n$ in time $T > 0$. The first equation is bilinear. The reachability of this system follows from a beautiful argument by Brockett (2012) that we now sketch. Recall that a bilinear control system on a Lie group is reachable provided the system is accessible and there exists a constant control such that the system is weakly Poisson stable; see Lian, Wang and Fu (1994). Since (A,B) is reachable, there exists a feedback matrix K such that $L := A + BK$ has distinct eigenvalues that are integer multiples of $2\pi\sqrt{-1}$. Thus there exists a (constant) control K such that e^{tL} is periodic. This shows that $\dot{X}(t) = (A + BK(t))X(t)$ is weakly Poisson stable for a suitable constant input. So it remains to prove the accessibility of the system. To this end, we compute the system Lie algebra $\mathfrak{g}$. Thus $\mathfrak{g}$ contains A together with every square matrix whose image is contained in that of B. The Lie bracket of A and BK is $[A, BK] = ABK - BKA$. Hence, $\mathfrak{g}$ contains every matrix whose image space is contained in that of AB. By iterating this argument, we see that $\mathfrak{g}$ contains all matrices whose image space is contained in the image space of $(B, AB, \ldots, A^{nB})$. By the reachability of (A,B), this implies that $\mathfrak{g}$ contains all real $n \times n$ matrices. Therefore, the Lie algebra rank condition shows the accessibility of the bilinear system. ■

The preceding proof shows that open-loop control enables one only to control the mean value of a Gaussian state distribution. To control both the mean and variance of a Gaussian distribution, one needs to apply both open-loop and closed-loop controls in the bilinear affine form $u(t,x) = K(t)x + u(t)$. We conclude with the following straightforward consequence of Theorem 12.26.

Theorem 12.27. *Assume that (A,B) is reachable. The transport system*

$$\frac{\partial \rho(t,x)}{\partial t} = -tr\,(A+BK(t))\rho(t,x) - ((A+BK(t))x+Bu(t))^\top \nabla\rho(t,x), \qquad (12.39)$$

with independent controls $u(t)$ and $K(t)$, leaves the manifold $\mathscr{G}$ of Gaussian distributions invariant. Using the controlled flow (12.39), one can steer in finite time $T>0$ two Gaussian distributions g_{Q_1,μ_1} and g_{Q_2,μ_2} into each other. Thus (12.39) is reachable on $\mathscr{G}$.

The preceding result can be generalized in several directions. First, one might consider replacing the set of positive definite matrices by the positive cone in a Euclidean Jordan algebra. Instead of using state feedback $A+BK$, one could study the effects of output feedback $A+BKC$. Finally, one might consider networks of systems and try to establish controllability results for the mean and covariance of state vectors in such interconnected systems. We leave these problems for future research.

12.5 Exercises

1. Show that the discrete-time system $x_{t+1}(\theta) = A(\theta)x_t(\theta) + b(\theta)u_t$, with

$$A(\theta) = \begin{pmatrix} \theta & 1 \\ 0 & \theta \end{pmatrix}, \quad b(\theta) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

   is not ensemble reachable over $\mathbf{P} = [0,1]$.
2. Show that the discrete-time system $x_{t+1}(\theta) = A(\theta)x_t(\theta) + b(\theta)u_t$, with

$$A(\theta) = \begin{pmatrix} 0 & 1 \\ 0 & \theta \end{pmatrix}, \quad b(\theta) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

   is ensemble reachable over $\mathbf{P} = [0,1]$.
3. Let $\mathbf{P} = [0,1]$. Consider the infinite sequence of systems

$$A_k(\theta) = \begin{pmatrix} 0 & -\theta^2 - 1/k^2 \\ 0 & 2\theta \end{pmatrix}, \quad b_k(\theta) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

   a. Verify that $A_k(\theta)$ and $b_k(\theta)$ are uniformly ensemble reachable for each finite k, but the limiting system $A_\infty(\theta), b_\infty(\theta))$ is not uniformly ensemble reachable.
   b. Prove that the set of uniformly ensemble reachable SISO systems is neither open nor closed in the topology of uniform convergence.

4. Let $(A,b,c) \in \mathbb{R}^{n\times n} \times \mathbb{R}^n \times \mathbb{R}^{1\times n}$ be reachable and observable, $R > 0$ fixed, and assume that $A - \theta^2 bc$ has distinct eigenvalues for $\theta \in \mathbf{P} = [0,R]$. Let $x_*(\theta)$ be a continuous function on $\mathbf{P}$. Prove that there exists $T > 0$ and a piecewise constant input function $u : [0,T] \longrightarrow \mathbb{R}$ such that the Fourier transform of the unique solution $\psi(t,z) \in \mathbb{R}^n$ of the coupled system of PDEs
$$\frac{\partial \psi}{\partial t}(t,z) = (A + \frac{\partial^2}{\partial z^2} bc)\psi(t,z) + bu(t)\sqrt{\frac{2}{\pi}}\frac{sin(Rz)}{z}, \quad \psi(0,z) = 0,$$
satisfies
$$\max_{\theta \in \mathbf{P}} \|(\mathscr{F}\psi)(T,\theta) - x_*(\theta)\| < \varepsilon.$$

5. Prove the following formula for the L^2-distance of two Gaussian distributions:
$$F(Q_1,Q_2,\mu_1,\mu_2) := \int_{\mathbb{R}^n} (g_{Q_1,\mu_1}(x) - g_{Q_2,\mu_2}(x))^2 dx = \frac{1}{\sqrt{(4\pi)^n \det Q_1}} +$$
$$\frac{1}{\sqrt{(4\pi)^n \det Q_2}} - \frac{2\exp\left(-\frac{1}{2}(\mu_1 - \mu_2)^\top (Q_1 + Q_2)^{-1}(\mu_1 - \mu_2)\right)}{\sqrt{(2\pi)^n \det(Q_1 + Q_2)}}.$$
Show that for fixed values of Q_1 and μ_1 the function $(Q,\mu) \mapsto F(Q_1,Q,\mu_1,\mu)$ is strictly convex and assumes its minimum value at $Q = Q_1, \mu = \mu_1$.

12.6 Notes and References

Open-loop control issues for ensembles of systems have been studied in order to design robust compensating pulse sequences in quantum control and NMR spectroscopy; see, for example, Li and Khaneja (2006). The main motivation here comes from the known difficulties of quantum mechanical observations, which makes feedback strategies difficult to implement. Rigorous results for the ensemble control of infinite-dimensional bilinear systems have been obtained by Beauchard, Coron and Rouchon (2010). Open-loop control is also of interest for understanding biological formation control tasks for flocks of systems; see, for example, Brockett (2010). Perhaps the best control strategies are neither pure open-loop nor feedback control, and a mixture of the two seems more promising. This point of view has been stressed in a number of recent papers on minimum attention control, for example, by Brockett (1997), Brockett (2008), and Brockett (2012). These general control issues of how to best combine open- and closed-loop control strategies certainly deserve further study.

A well-studied issue in controlling parameter-dependent systems is that of robustness, where the goal is to find input functions that achieve a desired control objective, irrespective of parameter variations in the system. This can be done either in closed-loop, using feedback transformations, or open-loop control strategies. A classical robust feedback control problem is the so-called blending problem, namely, the task of finding a dynamic output feedback controller that simultaneously stabilizes a one-parameter family of scalar transfer functions; see, for example, Khargonekar and Tannenbaum (1985); Tannenbaum (1980) for a solution to special cases via Nevanlinna–Pick interpolation. To date, no general solution to the blending problem for families of MIMO systems is known. Proceeding in a different direction, we mention the pioneering work by Kharitonov on robust stability. In fact, the paper by Kharitonov (1978) has inspired several researchers to find switching controllers for the stabilization of polytopic families of linear systems; see, for example, Ghosh (1985).

The approximation theorem of Mergelyan (1952) is a very natural generalization of the Stone–Weierstrass theorem, but known proofs as in Gaier (1987) are not constructive. Runge's approximation theorem is a weaker version of Mergelyan's theorem, where the approximating polynomials get replaced by rational functions. Error results for the polynomial approximations in Mergelyan's theorem are obtained by Saff and Totik (1989). A potentially interesting method for computing input sequences for ensemble control is provided by the Faber polynomials (we are grateful to Christian Lubich for suggesting this to us). Faber polynomials $p_j(z), j \in \mathbb{N}$, allow one to approximate analytic functions $f(z)$ in a complex domain K by a convergent series of the form $c_0 + \sum_{j=1}^{\infty} c_j p_j(z)$, where only the coefficients c_j depend on $f(z)$. Such polynomials exist if the complement of K in the extended complex plane is simply connected.

Corollary 12.4 characterizes ensemble reachability for a finite set of parameters. The proof shows that this statement, i.e., the characterization of reachability for the parallel connection of finitely many SISO systems, is equivalent to the Chinese remainder theorem or to Lagrange interpolation. Thus corresponding reachability results should follow for parallel connections of countably many SISO systems via interpolation results for analytic functions such as the Mittag–Leffler theorem. We refer to Helmke and Schönlein (2014) for a proof of Theorem 12.6. Condition (c) in Theorem 12.6, stating that all eigenvalues of $A(\theta)$ are simple, cannot be removed easily. Exercise 1 gives a counterexample. Proposition 12.12 and Exercises 1 and 3 are due to Scherlein (2014).

It has been shown that, by restricting the set of parameters in (12.1) to a finite subset, a parametric family of systems is equivalent to the parallel connection of linear systems. This can be generalized as follows. Consider a mixed differential and integral Volterra equation of the form

$$\frac{\partial}{\partial t} x(t,\theta) = A(\theta)x(t,\theta) + \int_{\mathbf{P}} K(\theta,\theta')x(t,\theta')d\theta' + B(\theta)u(t). \tag{12.40}$$

Here we allow for rather general classes of kernel functions $K(x,y)$. Note that if $\mathbf{P}$ is a compact group and the integral defines a convolution operator with respect to the Haar measure, then the class of spatially invariant systems studied by Bamieh, Paganini and Dahleh (2002) is obtained. Moreover, if one replaces the integral term with a Riemann sum and restricts oneself to the finite subset of N sampling points θ_i, then one obtains

$$\frac{\partial}{\partial t}x(t,\theta_i) = A(\theta_i)x(t,\theta_i) + \sum_{j=1}^{N} K(\theta_i,\theta_j)x(t,\theta_j) + B(\theta_i)u(t),$$

i.e., one obtains the general equations for a linear network of systems (9.5) studied in Chapter 9. This shows that spatiotemporal systems of the form (12.40) are the infinite-network counterparts of the finite system interconnections studied in Chapter 9. For general existence and uniqueness results for integral equations of the Volterra type we refer the reader to Väth (2000). We are not aware of systematic studies of the reachability or observability properties of systems of the form 12.40.

Theorem 12.19 and generalizations to exact reachability are due to Curtain, Iftime and Zwart (2009). The textbook by Curtain and Zwart (1995), Section 4.2, provides simple sufficient conditions for approximate reachability in a Hilbert space. These conditions require knowledge of a Riesz basis of eigenvectors and therefore do not apply to multiplication operators, which have a continuous spectrum. Multiplication operators on spaces of L^2-functions are not compact; the spectral approximation properties of such operators using finite-dimensional operators therefore become a nontrivial issue. We refer to Morrison (1995) for a nice presentation of results and examples in this direction.

The reachability properties of systems in Hardy spaces of analytic functions were derived by Fuhrmann (1972) and depend on Beurling's characterization of shift-invariant subspaces. We refer the reader to Fuhrmann (2012) for a discussion of SISO systems in a Hardy-space context that is close to the spirit of this book. A characterization of cyclic vectors for the backward shift is due to Douglas, Shapiro and Shields (1970). Composition operators on H^2 provide interesting examples of infinite-dimensional dynamical systems, and indeed of control systems. Littlewood's subordination principle, see, for example, Shapiro (1993), asserts that every composition operator $C_\phi(f) = f \circ \phi$ by an analytic function $\phi : \mathbb{D} \longrightarrow \mathbb{D}$, with $\phi(0) = 0$, takes the Hardy space H^2 into itself. This implies a generalization of the situation studied in Section 11.3 to infinite homogeneous networks, i.e., that the network transfer function $\mathcal{N}_g(z) = \mathcal{N}(h(z))$ of a homogeneous network with interconnection transfer function $\mathcal{N} \in H^2$ is always in H^2 provided $h(z) = 1/g(z)$ is an inner function with $h(0) = 0$. A nontrivial control system on spaces of univalent analytic functions is defined by Löwner's equation on the unit disc $\mathbb{D}$

$$\frac{\partial w(t,z)}{\partial t} = -\frac{e^{\sqrt{-1}u(t)} + w(t,z)}{e^{\sqrt{-1}u(t)} - w(t,z)}w(t,z), \quad w(0,z) = z.$$

Here the complex parameter z varies in the open unit disc. A generalization of this system on suitable matrix balls and its reachability properties would be interesting to study. For a study of the cyclicity and hyper cyclicity of composition operators, and further connections to universality in a function-theoretic context, we refer the reader to Grosse-Erdmann (1999). It seems that the connection to questions of reachability and observability has been overlooked within this circle of ideas.

A generalization of the Liouville equation (12.35) that has been frequently studied is the Fokker–Planck equation

$$\frac{\partial}{\partial t}\rho(x,t) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\left(a_{ij}(x,t,u)\rho(x,t)\right) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(b_i(x,t,u)\rho(x,t)\right).$$

It is well known that the Fokker–Planck equation describes the evolution of probability density functions that are propagated by stochastic differential equations; see, for example, Hazewinkel and Willems (1981) for a collection of articles on the subject. Thus, similar to the Liouville equation, the Fokker–Planck equation is a natural object of study in ensemble control. We refer the reader to Jordan, Kinderlehrer and Otto (1998) for a demonstration of the connection with steepest descent flows on spaces of probability measures with respect to the Wasserstein metric and to Blaquiere (1992) and Poretta (2014) for reachability results on the Fokker–Planck equation.