



Controllability Properties and Invariance Pressure for Linear Discrete-Time Systems

Fritz Colonius¹ · João A. N. Cossich² · Alexandre J. Santana²

Received: 3 September 2019 / Revised: 2 February 2021 / Accepted: 11 February 2021 /
Published online: 1 March 2021
© The Author(s) 2021

Abstract

For linear control systems in discrete time controllability properties are characterized. In particular, a unique control set with nonvoid interior exists and it is bounded in the hyperbolic case. Then a formula for the invariance pressure of this control set is proved.

Keywords Controllability · Control sets · Invariance pressure · Invariance entropy · Discrete-time control systems

Mathematics Subject Classification 93B05 · 37B40 · 94A17

1 Introduction

Invariance pressure for subsets of the state space generalizes invariance entropy of deterministic control systems by adding potentials on the control range. We consider control systems in discrete time of the form

$$x_{k+1} = F(x_k, u_k), k \in \mathbb{N}_0 = \{0, 1, \dots\},$$

where $F : M \times U \rightarrow M$ is smooth for a smooth manifold M and a compact control range $U \subset \mathbb{R}^m$. The invariance entropy $h_{inv}(K, Q)$ determines the average data rate needed to keep the system in Q (forward in time) when it starts in $K \subset Q$. Basic references for invariance entropy are Nair et al. [12] and the monograph Kawan [10], where also the relation to minimal data rates is explained. With some analogy to classical constructions for dynamical systems, invariance pressure adds continuous functions $f : U \rightarrow \mathbb{R}$ called potentials giving a weight to the control values.

We have announced some results of the present paper in “Invariance pressure for linear discrete-time systems”, Proceedings of the 2019 IEEE Information Theory Workshop (IEEE ITW 2019), Visby, Sweden, 24–26 Aug. 2019.

✉ Fritz Colonius
fritz.colonius@math.uni-augsburg.de

¹ Institut für Mathematik, Universität Augsburg, Augsburg, Germany

² Departamento de Matemática, Universidade Estadual de Maringá, Maringá, Brazil

For continuous-time systems, invariance entropy of hyperbolic control sets has been analyzed in Kawan [9] and Kawan and Da Silva [5]. Kawan and Da Silva [11] and [6] analyze invariance entropy of partially hyperbolic controlled invariant sets and chain control sets. Huang and Zhong [8] show dimension-like characterizations of invariance entropy. Measure-theoretic versions of invariance entropy have been considered in Colonius [4] and Wang et al. [15]. Invariance pressure has been analyzed in Colonius et al. [1–3]. In Zhong and Huang [18] it is shown that several generalized notions of invariance pressure fit into the dimension-theoretic framework due to Pesin.

The main results of the present paper are given for linear control systems $x_{k+1} = Ax_k + Bu_k$ with an invertible matrix A and control values u_k in a compact neighborhood U of the origin in \mathbb{R}^m . It is shown that a unique control set D with nonvoid interior exists if and only if the system without control constraints is controllable (i.e., the pair (A, B) is controllable), and D is bounded if and only if A is hyperbolic. In this case a formula for the invariance pressure of compact subsets K in D is presented.

The contents of this paper are as follows: Sect. 2 collects general properties of control sets for nonlinear discrete-time systems. Section 3 characterizes controllability properties of linear discrete-time systems with control constraints and Sect. 4 shows that here a unique control set with nonvoid interior exists and that it is bounded if and only if the uncontrolled system is hyperbolic. Section 5 introduces invariance entropy and as a generalization total invariance pressure where potentials on the product of the state space and the control range are allowed. For linear systems, Sect. 6 first derives an upper bound for the total invariance pressure and a lower bound for the invariance pressure. Combined they yield a formula for the invariance pressure in the hyperbolic case.

2 Control Sets for Nonlinear Systems

In this section we introduce some notation and prove several properties of control sets with nonvoid interior for nonlinear discrete-time systems. They are analogous to properties of systems in continuous time, however, the statements are a bit more involved, since one has to consider in addition to the interior of control sets their transitivity sets. A discussion of various slightly differing versions in the literature is contained in Colonius [4, Section 5].

We consider control systems of the form

$$x_{k+1} = F(x_k, u_k), k \in \mathbb{N}_0, \quad (1)$$

on a C^∞ -manifold M of dimension d endowed with a corresponding metric. For an initial value $x_0 \in M$ at time $k = 0$ and control $u = (u_k)_{k \geq 0} \in \mathcal{U} := U^{\mathbb{N}_0}$ we denote the solutions by $\varphi(k, x_0, u)$, $k \in \mathbb{N}_0$. Assume that the set of control values $U \subset \mathbb{R}^m$ is nonvoid and satisfies $U \subset \overline{\text{int}U}$. Let \tilde{U} be an open set containing \bar{U} and suppose that the map $F : M \times \tilde{U} \rightarrow M$ is a C^∞ -map.

Definition 1 For $x \in M$ and $k \in \mathbb{N}$ the reachable set $\mathbf{R}_k(x)$ and the controllable set $\mathbf{C}_k(x)$ are

$$\begin{aligned} \mathbf{R}_k(x) &:= \{y \in M \mid \exists u \in \mathcal{U} : y = \varphi(k, x, u)\}, \\ \mathbf{C}_k(x) &:= \{y \in M \mid \exists u \in \mathcal{U} : \varphi(k, y, u) = x\}, \end{aligned}$$

resp., and $\mathbf{R}(x)$ and $\mathbf{C}(x)$ are the respective unions over all $k \in \mathbb{N}$. The system is called accessible in x if

$$\text{int}\mathbf{R}(x) \neq \emptyset \text{ and } \text{int}\mathbf{C}(x) \neq \emptyset. \tag{2}$$

Accessibility in x certainly holds if

$$\text{int}F(x, U) \neq \emptyset \text{ and } \text{int}\{y \in M \mid x \in F(y, U)\} \neq \emptyset.$$

Next we specify maximal subsets of complete approximate controllability.

Definition 2 For system of the form (1) a nonvoid subset $D \subset M$ is called a control set if it is maximal with (i) $D \subset \overline{\mathbf{R}(x)}$ for all $x \in D$, (ii) for every $x \in D$ there is $u \in U$ with $\varphi(k, x, u) \in D$ for all $k \in \mathbb{N}$. The transitivity set D_0 of D is $D_0 := \{z \in D \mid z \in \text{int}\mathbf{C}(z)\}$.

We define for $k \geq 1$ a C^∞ -map

$$G_k : M \times U^k \rightarrow M, G_k(x, u) := \varphi(k, x, u).$$

Following Wirth [17] we say that a pair $(x, u) \in M \times \text{int}U^k$ is regular if $\text{rank} \frac{\partial G_k}{\partial u}(x, u) = d$ (clearly, this implies $mk \geq d$). For $x \in M$ and $k \in \mathbb{N}$ the regular reachable set and the regular controllable set at time k are

$$\begin{aligned} \hat{\mathbf{R}}_k(x) &:= \{\varphi(k, x, u) \mid (x, u) \text{ is regular}\}, \\ \hat{\mathbf{C}}_k(x) &:= \{y \in M \mid x = \varphi(k, y, u) \text{ with } (y, u) \text{ regular}\}, \end{aligned}$$

resp., and the regular reachable set $\hat{\mathbf{R}}(x)$ and controllable set $\hat{\mathbf{C}}(x)$ are given by the respective union over all $k \in \mathbb{N}$. It is clear that $\hat{\mathbf{R}}(x)$ and $\hat{\mathbf{C}}(x)$ are open for every x .

Accessibility condition (2) implies that there is $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$ one has $\text{int}\mathbf{R}_k(x) \neq \emptyset$ and

$$\mathbf{R}_k(x) \subset \overline{\{\varphi(k, x, u) \in \text{int}\mathbf{R}_k(x) \mid u \in \text{int}U^k\}}.$$

By Sard’s Theorem the set of points $\varphi(k, x, u) \in \mathbf{R}_k(x)$ such that (x, u) is not regular has Lebesgue measure zero.

Proposition 3 Assume that accessibility condition (2) holds for all $x \in M$. Then for every control set D with nonvoid interior the transitivity set D_0 is nonvoid and dense in $\text{int}D$.

Proof For $x \in \text{int}D$ there is $k_0 \in \mathbb{N}$ such that the reachable set $\mathbf{R}_k(x)$ at time k has nonvoid interior for all $k \geq k_0$. There is $k \geq k_0$ with $\mathbf{R}_k(x) \cap \text{int}D \neq \emptyset$, hence we may assume that there is $y := \varphi(k, x, u) \in \text{int}\mathbf{R}_k(x) \cap \text{int}D$. Then, by Sard’s Theorem, it follows that there is a point $y = \varphi(k, x, u) \in \text{int}D$ with some regular (x, u) , i.e., $y \in \text{int}D \cap \hat{\mathbf{R}}_k(x)$. Then $x \in \text{int}\mathbf{C}(y)$. Let $V \subset \text{int}\mathbf{C}(y)$ be a neighborhood of x . Since $x \in \text{int}D$ and $D \subset \overline{\mathbf{R}(y)}$, there is $z \in V \cap \mathbf{R}(y) \subset D$ and thus $y \in \mathbf{C}(z)$. By construction, the point $z \in D$ satisfies $z \in \text{int}\mathbf{C}(y) \subset \text{int}\mathbf{C}(z)$, hence it is in the transitivity set of D and D_0 is dense in $\text{int}D$. \square

Remark 4 In the general context of semigroups of continuous maps (and with slightly different notation), Patrão and San Martin [13, Propositions 4.8 and 4.10] show that the transitivity set D_0 is dense in a control set D with nonvoid interior provided that $D_0 \neq \emptyset$.

We note the following further results for control sets.

Proposition 5 Assume that D is a control set for a control system which is accessible for all $x \in M$. Then its transitivity set D_0 satisfies $D_0 \subset \mathbf{R}(x)$ for all $x \in D$.

Proof Let $x \in D$ and $x_0 \in D_0$. By approximate controllability of D and $x_0 \in \text{int}\mathbf{C}(x_0)$, there are $k \geq 1$ and $u \in \mathcal{U}$ with $\varphi(k, x, u) \in \text{int}\mathbf{C}(x_0)$. Hence there are $l \geq 1$ and $v \in \mathcal{U}$ such that $\varphi(l, \varphi(k, x, u), v) = x_0$. Therefore $x \in \mathbf{C}(x_0)$, that is, $x_0 \in \mathbf{R}(x)$. \square

Proposition 6 Assume that D is a control set with nonvoid interior of a control system, which is accessible for all $x \in M$. Then the transitivity set D_0 of D is nonvoid and

$$D = \overline{\mathbf{R}(x_0)} \cap \mathbf{C}(x_0) \text{ for all } x_0 \in D_0,$$

in particular, the set D is measurable.

Proof By Proposition 3 the transitivity set D_0 is nonvoid. Let $x_0 \in D_0$. Note that $D \subset \overline{\mathbf{R}(x_0)}$ by the definition of control sets. For every $x \in D$, Proposition 5 shows that $x_0 \in \mathbf{R}(x)$, that is $x \in \mathbf{C}(x_0)$. Hence $D \subset D' := \overline{\mathbf{R}(x_0)} \cap \mathbf{C}(x_0)$. It is not difficult to see that the set D' is a set of approximate controllability with nonvoid interior. It follows that D' is contained in a maximal set D'' of approximate controllability with nonvoid interior, which by Kawan [10, Proposition 1.20] is a control set. By the maximality property of control sets and $D \subset D''$, it follows that $D = D' = D''$, which concludes the proof. \square

The following proposition shows that a trajectory starting in the interior of a control set D and remaining in it up to a positive time must actually remain in the interior of D .

Proposition 7 Assume that the maps $F(\cdot, u)$ are local diffeomorphisms on M for all $u \in U$. Let x be in the interior of a control set D and suppose that for some $\tau \in \mathbb{N}$ and $u \in \mathcal{U}$ one has $\varphi(k, x, u) \in D$, $k \in \{1, \dots, \tau\}$. Then $\varphi(k, x, u) \in \text{int}D$, $k \in \{1, \dots, \tau\}$.

Proof Suppose that $y := \varphi(k, x, u) \in D \cap \partial D$ for some $k \in \{1, \dots, \tau\}$. By the assumption on the maps $F(\cdot, u)$ and $x \in \text{int}D$, there is a neighborhood $N_0(y)$ of y with $N_0(y) = \varphi(k, N(x), u)$ for a neighborhood $N(x) \subset D$ of x . Since $y \in D$, there are a control $v \in \mathcal{U}$ and $k_0 \in \mathbb{N}$ with $\varphi(k_0, y, v) \in \text{int}D$. Then there is a neighborhood $N_1(y)$ with $\varphi(k_0, N_1(y), v) \subset \text{int}D$. By the maximality property of control sets it follows that the neighborhood $N_0(y) \cap N_1(y)$ of y is contained in D , contradicting $y \in \partial D$. \square

3 Controllability Properties of Linear Systems

Next we consider linear control systems in \mathbb{K}^d , $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, of the form

$$x_{k+1} = Ax_k + Bu_k, \quad u_k \in U \subset \mathbb{K}^m, \quad (3)$$

where $A \in Gl(d, \mathbb{K})$ and $B \in \mathbb{K}^{d \times m}$ and the control range U is a compact convex neighborhood of $0 \in \mathbb{K}^m$ with $U = \overline{\text{int}U}$.

For initial value $x \in \mathbb{K}^d$ and control $u \in \mathcal{U} = U^{\mathbb{N}_0}$ the solutions of (3) are given by

$$\varphi(k, x, u) = A^k x + \sum_{i=0}^{k-1} A^{k-1-i} Bu_i, \quad k \in \mathbb{N}_0.$$

Where convenient, we also use the notation $\varphi_{k,u} := \varphi(k, \cdot, u) : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Note the following observation.

Proposition 8 For $x \in \mathbb{K}^d$ the reachable set $\mathbf{R}_k(x)$ at time k ,

$$\mathbf{R}_k(x) = \{y \in \mathbb{K}^d \mid \exists u \in \mathcal{U} \text{ with } \varphi(k, x, u) = y\}$$

is compact and convex.

Proof Convexity follows from the convexity of U . Since $U \subset \mathbb{K}^m$ is compact, there is $M > 0$ such that $\|u\| \leq M$, for all $u \in U$. Then, if $y = \varphi(k, x, u) \in \mathbf{R}_k(x)$, $u = (u_i) \in U^k$, we get

$$\|y\| \leq \|A^k x\| + \sum_{i=0}^{k-1} \|A^{k-1-i} B u_i\| \leq \|A\|^k \|x\| + M \sum_{i=0}^{k-1} \|A\|^{k-1-i} \|B\| < \infty,$$

hence $\mathbf{R}_k(x)$ is bounded. In order to show that $\mathbf{R}_k(x)$ is closed, consider a sequence $y_n = \varphi(k, x, u^n)$ in $\mathbf{R}_k(x)$ such that $y_n \rightarrow y \in \mathbb{K}^d$ and $u^n \in U^k$. By compactness of U , we have that U^k is compact, hence there is a subsequence converging to some $u \in U^k$. Therefore $y = \varphi(k, x, u) \in \mathbf{R}_k(x)$ by continuity. \square

Proposition 9 For all $k, l \in \mathbb{N}$ we have

$$\mathbf{R}_k(0) + A^k \mathbf{R}_l(0) = \mathbf{R}_{l+k}(0) \text{ and } \text{int} \mathbf{R}_k(0) + A^k \mathbf{R}_l(0) \subset \text{int} \mathbf{R}_{k+l}(0).$$

Proof Let $x_1 \in \mathbf{R}_k(0)$ and $x_2 \in \mathbf{R}_l(0)$. Then there are $u, v \in \mathcal{U}$ such that

$$x_1 = \sum_{i=0}^{k-1} A^{k-1-i} B u_i \text{ and } x_2 = \sum_{i=0}^{l-1} A^{l-1-i} B v_i.$$

Define

$$w_i = \begin{cases} v_i, & \text{if } 0 \leq i \leq l-1 \\ u_{i-l}, & \text{if } l \leq i \leq k+l-1 \end{cases}.$$

Then

$$\begin{aligned} \varphi(k+l, 0, w) &= \sum_{i=0}^{k+l-1} A^{k+l-1-i} B w_i = \sum_{i=0}^{l-1} A^{k+l-1-i} B w_i + \sum_{i=l}^{k+l-1} A^{k+l-1-i} B w_i \\ &= A^k \sum_{i=0}^{l-1} A^{l-1-i} B v_i + \sum_{i=0}^{k-1} A^{k-1-i} B u_i = A^k x_2 + x_1. \end{aligned}$$

Hence $x_1 + A^k x_2 = \varphi(k+l, 0, w) \in \mathbf{R}_{l+k}(0)$. The converse inclusion follows by reversing these steps. The second assertion follows since the set on left hand side is open. \square

Define the time reversed counterpart of system (3) by

$$x_{k+1} = A^{-1} x_k - A^{-1} B u_k, \quad u_k \in U \subset \mathbb{K}^m. \tag{4}$$

The reachable and controllable sets from the origin at time k for this system are denoted by $\mathbf{R}_k^-(0)$ and $\mathbf{C}_k^-(0)$, respectively.

Proposition 10 The reachable and controllable sets for system (3) and the time reversed system (4) satisfy for all $k \in \mathbb{N}$

$$\mathbf{R}_k(0) = \mathbf{C}_k^-(0) \text{ and } \mathbf{C}_k(0) = \mathbf{R}_k^-(0).$$

Proof Note that $x \in \mathbf{C}_k(0)$ if and only if there is $u \in \mathcal{U}$ with

$$A^k x + \sum_{i=0}^{k-1} A^{k-1-i} B u_i = 0, \text{ i.e., } x = - \sum_{i=0}^{k-1} A^{-1-i} B u_i.$$

For any $u \in U^k$, we define $v_j = u_{k-1-j}$, $0 \leq j \leq k - 1$. Then

$$\begin{aligned} x &= - \sum_{i=0}^{k-1} A^{-1-i} B u_i = - \sum_{j=0}^{k-1} A^{-1-(k-1-j)} B u_{k-1-j} = - \sum_{j=0}^{k-1} (A^{-1})^{k-j} B v_j \\ &= - \sum_{j=0}^{k-1} (A^{-1})^{k-1-j} A^{-1} B v_j = \sum_{j=0}^{k-1} (A^{-1})^{k-1-j} (-A^{-1} B) v_j. \end{aligned}$$

Hence we conclude that $x \in \mathbf{C}_k(0)$ if and only if there exists a control $v \in U^k$ such that $x = \varphi^-(k, 0, v)$, where φ^- is the solution of (4). This proves that $\mathbf{C}_k(0) = \mathbf{R}_k^-(0)$. The other equality follows analogously. \square

Lemma 11 *If (A, B) is controllable, there is $\delta > 0$ such that the ball $B_\delta(0)$ satisfies $B_\delta(0) \subset \text{int}\mathbf{R}_{d-1}(0)$. Furthermore, $\mathbf{R}_n(0) \subset \mathbf{R}_m(0)$ for $m \geq n$.*

Proof Since the control range is a neighborhood of 0, controllability implies that there is $\delta > 0$ with $B_\delta(0) \subset \text{int}\mathbf{R}_{d-1}(0)$. The second assertion follows since 0 is an equilibrium for $u = 0$. \square

Proposition 12 *If (A, B) is controllable, the reachable set of system (3) satisfies $\overline{\mathbf{R}(0)} = \text{int}\mathbf{R}(0)$.*

Proof The inclusion $\overline{\text{int}\mathbf{R}(0)} \subset \overline{\mathbf{R}(0)}$ holds trivially. For the converse we first show that $\mathbf{R}(y) \subset \text{int}\mathbf{R}(0)$ for $y \in \text{int}\mathbf{R}(0)$. In fact, let there exists a neighborhood V_y of y such that $V_y \subset \mathbf{R}(0)$. Given $z \in \mathbf{R}(y)$, there are $k \in \mathbb{N}$ and $u \in \mathcal{U}$ such that $z = \varphi(k, y, u)$. Since $A \in Gl(d, \mathbb{R})$, the map $\varphi_{k,u}$ is a diffeomorphism and we have that $\varphi_{k,u}(V_y)$ is a neighborhood of z and clearly $\varphi_{k,u}(\mathbf{R}(0)) \subset \mathbf{R}(0)$. So $z \in \varphi_{k,u}(V_y) \subset \mathbf{R}(0)$, which shows that $z \in \text{int}\mathbf{R}(0)$.

Now, let $x \in \overline{\mathbf{R}(0)}$ and V a neighborhood of x . There is $y \in \mathbf{R}(0)$ such that $y \in V$, so there are $k \in \mathbb{N}$ and $u \in \mathcal{U}$ such that $y = \varphi(k, 0, u)$. Since $0 \in \text{int}\mathbf{R}(0)$ there exists a neighborhood W of 0 such that $W \subset \text{int}\mathbf{R}(0)$ and $\varphi_{k,u}(W) \subset V$ by continuity of $\varphi_{k,u}$. For $z \in W$ the arguments above show that $\mathbf{R}(z) \subset \text{int}\mathbf{R}(0)$ and it follows that

$$\varphi(k, z, u) \in V \cap \mathbf{R}(z) \subset V \cap \text{int}\mathbf{R}(0)$$

and hence $x \in \overline{\text{int}\mathbf{R}(0)}$. \square

We will need the following lemmas.

Lemma 13 *For every $\lambda \in \mathbb{C}$ there are $n_k \rightarrow \infty$ such that $\frac{\lambda^{n_k}}{|\lambda|^{n_k}} \rightarrow 1$, and, in particular,*

$$\frac{\text{Im}(\lambda^{n_k})}{\text{Re}(\lambda^{n_k})} \rightarrow 0 \text{ for } k \rightarrow \infty.$$

Proof There is $\theta \in [0, 2\pi)$ with $\lambda = |\lambda|(\cos \theta + \iota \sin \theta)$, hence

$$\lambda^n = |\lambda|^n (\cos(n\theta) + \iota \sin(n\theta)).$$

If $\theta \in 2\pi\mathbb{Q}$, there are $n, N \in \mathbb{N}$ with $n\theta = N2\pi$, hence $\lambda^n = |\lambda|^n \cos(N2\pi) = |\lambda|^n$. Else, there are $n_k \rightarrow \infty$ such that modulo 2π one has $n_k\theta \rightarrow 0$. This implies $\cos(n_k\theta) \rightarrow 1$ and $\sin(n_k\theta) \rightarrow 0$, hence

$$\frac{\lambda^{n_k}}{|\lambda|^{n_k}} = \cos(n_k\theta) + \iota \sin(n_k\theta) \rightarrow 1.$$

This implies

$$\frac{\text{Im}(\lambda^{n_k})}{\text{Re}(\lambda^{n_k})} = \frac{\text{Im}\left(\frac{\lambda^{n_k}}{|\lambda|^{n_k}}\right)}{\text{Re}\left(\frac{\lambda^{n_k}}{|\lambda|^{n_k}}\right)} = \frac{\sin(n_k\theta)}{\cos(n_k\theta)} \rightarrow 0.$$

□

The next lemma states a property of convex sets.

Lemma 14 *If C is an open convex subset of \mathbb{K}^n and $Y \subset C$ a subspace, then $C = C + Y$.*

The following theorem describes the general structure of reachable and controllable sets. It is analogous to a well known property of linear systems in continuous time, cf. Sontag [14, Section 3.6] and Hinrichsen and Pritchard [7, Theorem 6.2.15]; the proof for discrete-time systems, however, is more involved. Recall that the state space \mathbb{K}^d can be decomposed with respect to A into the direct sum of the stable subspace E^s , the center space E^c and the unstable subspace E^u which are the direct sums of all generalized (real) eigenspaces for the eigenvalues λ of A with $|\lambda| < 1$, $|\lambda| = 1$ and $|\lambda| > 1$, respectively. Furthermore, we let $E^{uc} := E^u \oplus E^c$ and $E^{sc} := E^s \oplus E^c$.

Theorem 15 *Consider the control system given by (3) and suppose that the system without control restriction is controllable.*

- (i) *There exists a compact and convex set $K \subset E^s \subset \mathbb{K}^d$ with nonvoid interior with respect to E^s such that $\overline{\mathbf{R}(0)} = K + E^{uc}$. Moreover $0 \in K$ and $E^{uc} \subset \text{int}\mathbf{R}(0)$.*
- (ii) *There exists a compact and convex set $F \subset E^u \subset \mathbb{K}^d$ with nonvoid interior with respect to E^u such that $\overline{\mathbf{C}(0)} = F + E^{sc}$. Moreover $0 \in F$ and $E^{sc} \subset \text{int}\mathbf{C}(0)$.*

Proof We will first prove the result for $\mathbb{K} = \mathbb{C}$.

- (i) In the first step, we will show that $E^{uc} \subset \text{int}\mathbf{R}(0)$. As $\mathbf{R}(0)$ is convex, its interior is convex too. Therefore it suffices to prove that the generalized eigenspaces for eigenvalues with absolute value greater than or equal to 1 are contained in $\text{int}\mathbf{R}(0)$. Fix an eigenvalue λ of A with $|\lambda| \geq 1$ and let $E_q(\lambda) = \ker(A - \lambda I)^q$, $q \in \mathbb{N}_0$. It suffices to show that $E_q(\lambda) \subset \text{int}\mathbf{R}(0)$ for all q .

We prove the statement by induction on q , the case $q = 0$ being trivial since $E_q(\lambda) = \{0\} \subset \text{int}\mathbf{R}(0)$. So assume that $E_{q-1}(\lambda) \subset \text{int}\mathbf{R}(0)$ and take any $w \in E_q(\lambda)$. We must show that $w \in \text{int}\mathbf{R}(0)$. By Lemma 11 there is $\delta > 0$ such that $aw \in \text{int}\mathbf{R}_{d-1}(0)$ for all $a \in \mathbb{C}$ with $|a| < \delta$.

Note that for all $|a| < \delta$ and all $n \geq 1$

$$\begin{aligned} A^n aw &= (A - \lambda I + \lambda I)^n aw = \sum_{j=0}^n \binom{n}{j} (A - \lambda I)^{n-j} \lambda^j aw \\ &= \lambda^n aw + \sum_{j=0}^{n-1} \binom{n}{j} (A - \lambda I)^{n-j} \lambda^j aw. \end{aligned}$$

Since $aw \in E_q(\lambda)$, it follows that $(A - \lambda I)^j aw \in E_{q-1}(\lambda)$ for all $i \geq 1$, hence $z(n) := \sum_{j=0}^{n-1} \binom{n}{j} (A - \lambda I)^{n-j} \lambda^j aw \in E_{q-1}(\lambda)$, $n \geq 1$. Using $aw \in \text{int}\mathbf{R}_{d-1}(0)$ Lemmas 11 and 14 imply for $n \geq 1$

$$\lambda^n aw = A^n aw - z(n) \in A^n aw + E_{q-1}(\lambda) \subset \text{int}\mathbf{R}_{n+d-1}(0) + E_{q-1}(\lambda) \subset \text{int}\mathbf{R}(0) \tag{5}$$

We write

$$a = \alpha + \iota\beta \text{ and } \lambda^n = x_n + \iota y_n$$

with $\alpha, \beta \in \mathbb{R}$ and $x_n, y_n \in \mathbb{R}^d$.

Claim: There are a sequence $(n_k)_{k \in \mathbb{N}}$ with $n_k \rightarrow \infty$ and $a_{n_k} \in \mathbb{C}$ with $|a_{n_k}| < \delta$ such that $\lambda^{n_k} a_{n_k} \in \mathbb{R}$.

In fact, we have

$$\lambda^n a = (x_n + \iota y_n)(\alpha + \iota\beta) = x_n\alpha - y_n\beta + \iota(x_n\beta + y_n\alpha) \in \mathbb{R},$$

if and only if $x_n\beta + y_n\alpha = 0$.

Case (a): If $x_n = 0$, one may choose $\alpha_n := 0$ and gets $\lambda^n a_n = -y_n\beta_n \in \mathbb{R}$ for $\beta_n = \frac{\delta}{2}$ with $|a_n| = |\beta_n| = \frac{\delta}{2}$.

Case (b): Otherwise $\lambda^n a \in \mathbb{R}$ if and only if

$$\beta = -\alpha \frac{y_n}{x_n} = -\alpha \frac{\text{Im}(\lambda^n)}{\text{Re}(\lambda^n)}.$$

According to Lemma 13 there are $n_k \in \mathbb{N}$, arbitrarily large, such that with $\alpha_{n_k} := \frac{\delta}{2}$ and $\beta_{n_k} := -\alpha_{n_k} \frac{y_{n_k}}{x_{n_k}}$

$$|\beta_{n_k}| = \frac{\delta}{2} \left| \frac{\text{Im}(\lambda^{n_k})}{\text{Re}(\lambda^{n_k})} \right| < \frac{\delta}{2}.$$

It follows for $a_{n_k} := \alpha_{n_k} + \beta_{n_k}$ that

$$|a_{n_k}|^2 = \alpha_{n_k}^2 + \beta_{n_k}^2 < \frac{1}{4}\delta^2 + \frac{1}{4}\delta^2, \text{ and hence } |a_{n_k}| < \delta.$$

We have shown that with this choice of a_{n_k} we have $\lambda^{n_k} a_{n_k} \in \mathbb{R}$ and the **Claim** is proved. Furthermore in case (a), by $|\lambda| \geq 1$,

$$|\lambda^n a_n| = |\lambda|^n |a_n| \geq |a_n| = \frac{\delta}{2},$$

and in case (b)

$$|\lambda^{n_k} a_{n_k}| = |\lambda|^{n_k} |a_{n_k}| \geq |a_{n_k}| \geq |\alpha_{n_k}| = \frac{\delta}{2}.$$

Now choose $\ell \in \mathbb{N}$ with $\ell \geq 2/\delta$. Recall that all points $a_{n_k} w \in \text{int}\mathbf{R}_{d-1}(0)$. We may assume that $n_2 \geq n_1 + d - 1$, hence

$$A^{n_1} a_{n_1} w \in \text{int}\mathbf{R}_{n_1+d-1}(0) \subset \text{int}\mathbf{R}_{n_2}(0).$$

We may also assume that $n_3 - n_2 \geq n_2 + d - 1$, hence

$$A^{n_2} a_{n_2} w \in \text{int}\mathbf{R}_{n_2+d-1}(0) \subset \text{int}\mathbf{R}_{n_3-n_2}(0).$$

Thus Proposition 9 implies

$$A^{n_1} a_{n_1} w + A^{n_2} a_{n_2} w \in \text{int}\mathbf{R}_{n_2}(0) + A^{n_2} \mathbf{R}_{n_3-n_2}(0) \subset \text{int}\mathbf{R}_{n_3-n_2+n_2}(0) = \text{int}\mathbf{R}_{n_3}(0).$$

Proceeding in this way, we finally arrive at

$$\sum_{k=1}^{\ell} A^{n_k} a_{n_k} w \in \text{int}\mathbf{R}_{n_\ell}(0).$$

Thus we find with (5),

$$\sum_{k=1}^{\ell} \lambda^{nk} a_{nk} w = \sum_{k=1}^{\ell} [A^{nk} a_{nk} w - z(n_k)] \in \text{int}\mathbf{R}_{n_\ell}(0) + E_{q-1}(\lambda) \subset \text{int}\mathbf{R}(0).$$

If $\lambda^{nk} a_{nk} > 0$ for all $k \in \{1, \dots, \ell\}$, then (the real number)

$$\sum_{k=1}^{\ell} \lambda^{nk} a_{nk} > \ell \cdot \delta/2 \geq 1.$$

For the k with $\lambda^{nk} a_{nk} < 0$, replace a_{nk} by $-a_{nk}$, to get the same conclusion. This shows that w is a convex combination of the points 0 and $\sum_{k=1}^{\ell} \lambda^{nk} a_{nk} w$ in $\text{int}\mathbf{R}(0)$, thus convexity of this set implies $w \in \text{int}\mathbf{R}(0)$ completing the induction step $E_q(\lambda) \subset \text{int}\mathbf{R}(0)$. Hence we have shown that $E^{uc} \subset \text{int}\mathbf{R}(0)$.

It remains to construct a set K as in the assertion. Define $K_0 := \text{int}\mathbf{R}(0) \cap E^s$. Then it follows that

$$K_0 + E^{uc} = (\text{int}\mathbf{R}(0) \cap E^s) + E^{uc} \subset \text{int}\mathbf{R}(0) + E^{uc} \subset \text{int}\mathbf{R}(0).$$

For the converse inclusion, let $v \in \text{int}\mathbf{R}(0)$, then $v = x + y$ where $x \in E^s$ and $y \in E^{uc}$, hence by Lemma 14,

$$x = v - y \in \text{int}\mathbf{R}(0) + E^{uc} = \text{int}\mathbf{R}(0),$$

which shows that $x \in K_0$ and therefore $v \in K_0 + E^s$. This shows that

$$K_0 + E^{uc} = \text{int}\mathbf{R}(0). \tag{6}$$

In order to show that K_0 is bounded, consider the projection $\pi : \mathbb{C}^d = E^s \oplus E^{uc} \rightarrow E^s$ along E^{uc} . Since E^s and E^{uc} are A -invariant, π commutes with A and we have $\pi A^n = A^n \pi$, for all $n \in \mathbb{N}_0$. For each $x \in K_0 = \text{int}\mathbf{R}(0) \cap E^s$, there are $k \in \mathbb{N}$ and $u = (u_i) \in \mathcal{U}$ such that

$$x = \sum_{i=0}^{k-1} A^{k-1-i} B u_i.$$

Since $A|_{E^s}$ is a linear contraction, there exist constants $a \in (0, 1)$ and $c \geq 1$ such that $\|A^n x\| \leq ca^n \|x\|$ for all $n \in \mathbb{N}$ and $x \in E^s$. Since U is compact, there is $M > 0$ such that $\|\pi B u\| \leq M$, for all $u \in U$, so

$$x = \pi(x) = \pi \left(\sum_{i=0}^{k-1} A^{k-1-i} B u_i \right) = \sum_{i=0}^{k-1} \pi A^{k-1-i} B u_i = \sum_{i=0}^{k-1} A^{k-1-i} \pi B u_i,$$

hence

$$\|x\| \leq \sum_{i=0}^{k-1} \|A^{k-1-i} \pi B u_i\| \leq \sum_{i=0}^{k-1} \|A^{k-1-i}\| \|\pi B u_i\| \leq cM \sum_{i=0}^{k-1} a^{k-1-i} = cM \frac{1-a^k}{1-a}$$

showing that K_0 is bounded. As a consequence, $K := \overline{K_0} = \overline{\text{int}\mathbf{R}(0) \cap E^s}$ is a compact convex set which has nonvoid interior relative to E^s . Moreover, $K + E^{uc}$ is closed, because K is compact. Therefore it follows from Proposition 12 and (6) that

$$\overline{\mathbf{R}(0)} = \overline{\text{int}\mathbf{R}(0)} = \overline{K_0 + E^{uc}} = \overline{K + E^{uc}}.$$

(ii) Consider the time reversed system (4). Note that $\mathbb{C}^d = E_-^s \oplus E_-^c \oplus E_-^u$, where E_-^s, E_-^c and E_-^u are the sums of the generalized eigenspaces for the eigenvalues μ of A^{-1} with $|\mu| < 1, |\mu| = 1$ and $|\mu| > 1$, respectively. Now λ is an eigenvalue of A (note that $\lambda \neq 0$ since $A \in Gl(d, \mathbb{C})$), if and only if $\mu = \lambda^{-1}$ is an eigenvalue of A^{-1} . Hence we have $E_-^s = E^u, E_-^c = E^c$ and $E_-^u = E^s$. By (i) there exists a compact and convex set $F \subset \mathbb{C}^d$ which has nonvoid interior with respect to $E_-^s = E^u$ such that $\mathbf{R}^-(0) = F + E_-^{uc}, 0 \in F$ and $E_-^{uc} \subset \text{int}\mathbf{R}^-(0)$. By Proposition 10,

$$E^{sc} = E_-^{uc} \subset \text{int}\mathbf{R}^-(0) = \text{int}\mathbf{C}(0)$$

and

$$\overline{\mathbf{C}(0)} = F + E_-^{uc} = F + E^{sc}.$$

This completes the proof of the theorem for the case $\mathbb{K} = \mathbb{C}$.

It remains to prove the theorem for the case $\mathbb{K} = \mathbb{R}$. Note that if $A \in Gl(d, \mathbb{R})$, then $u - \iota v \in E^s, u, v \in \mathbb{R}^d$, implies $u + \iota v, v + \iota u \in E^s$ and a similar implication holds for E^{uc} . Hence

$$\begin{aligned} \text{Re } E^s &= E^s \cap \mathbb{R}^d, \text{Re } E^{uc} = E^{uc} \cap \mathbb{R}^d, \\ E^s &= \text{Re } E^s + \iota \text{Re } E^s, E^{uc} = \text{Re } E^{uc} \oplus \iota \text{Re } E^{uc} \end{aligned} \tag{7}$$

Let $U_{\mathbb{C}} := U + \iota U$ and apply the result above for $\mathbb{K} = \mathbb{C}$. Clearly (A, B) is controllable, when considered as a system with state space \mathbb{C}^d and $U_{\mathbb{C}}$ is a convex compact neighborhood of $0 \in \mathbb{C}^m$ with $U_{\mathbb{C}} \subset \overline{\text{int}U_{\mathbb{C}}}$.

Denote the reachable set from 0 of the real and complex system by $\mathbf{R}_{\mathbb{R}}$ and $\mathbf{R}_{\mathbb{C}}$, respectively. It follows from the complex version of the theorem that the compact convex set $K_{\mathbb{C}} := \overline{\text{int}(\mathbf{R}_{\mathbb{C}})} \cap E^s$ has non-empty interior relative to E^s and satisfies $\overline{\mathbf{R}_{\mathbb{C}}} = K_{\mathbb{C}} \cap E^{uc}$. Since every $u \in U_{\mathbb{C}}$ is of the form $u = v + \iota w$, where $v, w \in U$, and $\varphi(k, 0, u) = \varphi(k, 0, v) + \iota \varphi(k, 0, w), k \in \mathbb{N}$, we have

$$U_{\mathbb{C}} = U_{\mathbb{R}} + \iota U_{\mathbb{R}} \text{ and } \mathbf{R}_{\mathbb{C}} = \mathbf{R}_{\mathbb{R}} + \iota \mathbf{R}_{\mathbb{R}}. \tag{8}$$

It follows that

$$\mathbf{R}_{\mathbb{R}} = \text{Re } \mathbf{R}_{\mathbb{C}}, \text{int}\mathbf{R}_{\mathbb{R}} = \text{Re int}\mathbf{R}_{\mathbb{C}},$$

where the interior of $\mathbf{R}_{\mathbb{R}}$ is relative to \mathbb{R}^d and the interior of $\mathbf{R}_{\mathbb{C}}$ is relative to \mathbb{C}^d . Now, if $W, Z \subset \mathbb{C}^d$ are subsets of the form

$$W = W_1 + \iota W_2, Z = Z_1 + \iota Z_2,$$

where $W_1, W_2, Z_1, Z_2 \subset \mathbb{R}^d$ and $W \cap Z \neq \emptyset$, then $W \cap Z = (W_1 \cap Z_1) + \iota (W_2 \cap Z_2)$ and so $\text{Re}(W \cap Z) = \text{Re } W \cap \text{Re } Z$. Applying this equality to $W = \text{int}\mathbf{R}_{\mathbb{C}}$ and $Z = E^s$ we obtain from (8) and (7) that

$$K = \overline{(\text{Re}(\text{int}\mathbf{R}_{\mathbb{C}})) \cap \text{Re } E^s} = \overline{\text{Re}(\text{int}\mathbf{R}_{\mathbb{C}} \cap E^s)} = \text{Re } K_{\mathbb{C}}.$$

Hence K is a compact convex subset of \mathbb{R}^d , which has a non-empty interior relative to $\text{Re } E^s$. Using (8) for the second equality we get

$$\overline{\mathbf{R}_{\mathbb{R}}} = \overline{\text{Re } \mathbf{R}_{\mathbb{C}}} = \text{Re } \overline{\mathbf{R}_{\mathbb{C}}} = \text{Re}(K_{\mathbb{C}} + E^{u,s}) = K + \text{Re } E^{u,s}.$$

This concludes the proof. □

Next we present a necessary and sufficient condition for controllability in \mathbb{K}^d . This consequence of Theorem 15 illustrates that controllability only holds under very strong assumptions on the spectrum of the matrix A . In the next section, we will instead consider subsets of the state space where complete controllability holds, i.e., control sets. Recall that the system without control restriction is controllable in \mathbb{R}^d if and only if (A, B) is controllable.

Corollary 16 *Consider the discrete-time linear system given in (3).*

- (i) *The reachable set $\mathbf{R}(0) = \mathbb{K}^d$ if and only if (A, B) is controllable and A has no eigenvalues with absolute value less than 1.*
- (ii) *The controllable set $\mathbf{C}(0) = \mathbb{K}^d$ if and only if (A, B) is controllable and A has no eigenvalues with absolute value greater than 1.*
- (iii) *The system is controllable in \mathbb{K}^d if and only if (A, B) is controllable and all eigenvalues of A have absolute value equal to 1.*

Proof (i) If $\mathbf{R}(0) = \mathbb{K}^d$, then the pair (A, B) is controllable, since $\mathbf{R}(0)$ is contained in the image of Kalman’s matrix $[B \ AB \ \dots \ A^{d-1}B]$. Moreover, if there is an eigenvalue λ of A with $|\lambda| < 1$, then $E^s \neq \{0\}$ and E^u is a proper subset of \mathbb{K}^d . By Theorem 15 (ii), there is a nonvoid compact set $F \subset E^u$ such that $E^{sc} + F = \overline{\mathbf{R}(0)} = \mathbb{K}^d$, a contradiction.

Conversely, if (A, B) is controllable and all eigenvalues λ of A satisfy $|\lambda| \geq 1$, then by Theorem 15 (i) we have $\mathbb{K}^d = E^{uc} \subset \text{int}\mathbf{R}(0) \subset \mathbf{R}(0)$.

(ii) This follows analogously.

- (iii) This is a consequence of assertions (i) and (ii) observing that $\mathbf{R}(0) = \mathbf{C}(0) = \mathbb{K}^d$ holds if and only if for all $x, y \in \mathbb{K}^d$ there are a control $u \in \mathcal{U}$ and a time $k \in \mathbb{N}$ with $\varphi(k, x, u) = y$.

□

Remark 17 In the continuous-time case, a result analogous to Corollary 16 is given e.g. in Sontag [14, Section 3.6]. For the discrete-time case, we are not aware of a result in the literature covering Corollary 16. In the special case of two inputs (i.e., $m = 2$) the characterization of null-controllability in Corollary 16 (ii) is given in Wing and Desoer [16, Section V, Theorem 2].

4 Control Sets for Linear Systems

Next we analyze linear control systems in \mathbb{R}^d of the form

$$x_{k+1} = Ax_k + Bu_k, u_k \in U \subset \mathbb{R}^m \tag{9}$$

with $A \in Gl(d, \mathbb{R})$ and $B \in \mathbb{R}^{d \times m}$ and suppose that U is a convex compact neighborhood of $0 \in \mathbb{R}^m$ with $U = \text{int}U$. Recall that the system without control restrictions is controllable in \mathbb{R}^d if and only if $\text{rank}[B \ AB \ \dots \ A^{d-1}B] = d$, i.e., the pair (A, B) is controllable.

Theorem 18 *There exists a unique control set D with nonvoid interior of system (9) if and only if the system without control restriction is controllable in \mathbb{R}^d . In this case $0 \in D_0 \cap \text{int}D$.*

Proof The controllability condition for (A, B) is necessary for the existence of D , since it guarantees that accessibility condition (2) holds for all $x \in \mathbb{R}^d$ and, for the system without control constraints, the reachable and the null-controllable subspaces coincide with \mathbb{R}^d . Since $0 \in \text{int}U$, one verifies that for $k \geq d - 1$

$$0 \in \text{int}(\mathbf{C}_k(0)) \cap \text{int}(\mathbf{R}_k(0)) =: D'$$

Then every point $x \in D'$ can be steered to any other point $z \in D'$ (first steer x to the origin in time k and then the origin to z in time k) and $0 \in \text{int}(\mathbf{C}(0))$. As in the proof of Proposition 6 one finds that D' is contained in a control set D . Thus we have established the existence of a control set D with nonvoid interior, and $0 \in D_0 \cap \text{int}D$. It remains to show uniqueness.

Let $\tilde{D} \subset \mathbb{R}^d$ be an arbitrary control set with nonvoid interior. By Proposition 6 its transitivity set \tilde{D}_0 is nonvoid and for $x_0 \in \tilde{D}_0$

$$\tilde{D} = \overline{\mathbf{R}(x_0)} \cap \mathbf{C}(x_0).$$

By linearity, we have $\varphi(k, x_1, u) = x_2$ for $k \in \mathbb{N}$ and $x_1, x_2 \in \mathbb{R}^d$ implies $\varphi(k, \alpha x_1, \alpha u) = \alpha x_2$ for any $\alpha \in (0, 1]$. Here the control αu has values in U , since U is convex and $0 \in U$. This implies that $\alpha \tilde{D}$ is contained in some control set D^α and $\text{int}(\alpha \tilde{D})$ is contained in the interior of D^α . Now choose any $x \in \text{int} \tilde{D}$ and suppose, by way of contradiction, that

$$\alpha_0 := \inf\{\alpha \in (0, 1] \mid \forall \beta \in [\alpha, 1] : \beta x \in \tilde{D}\} > 0.$$

Then $\alpha_0 x \in \partial \tilde{D}$ and $\alpha_0 x \in \text{int}D^{\alpha_0}$. Therefore $\tilde{D} \cap \text{int}D^{\alpha_0} \neq \emptyset$, and it follows that $\tilde{D} = D^{\alpha_0}$ and $\alpha_0 x \in \text{int} \tilde{D}$. This is a contradiction and so $\alpha_0 = 0$. Choosing $\alpha > 0$ small enough such that $\alpha x \in D$, we obtain $\alpha x \in \tilde{D} \cap D \neq \emptyset$. Now it follows that $\tilde{D} = D$. \square

The following theorem gives a spectral characterization of boundedness of the control set. Recall that A is called hyperbolic if all eigenvalues λ of A satisfy $|\lambda| \neq 1$.

Theorem 19 *Assume that (A, B) is controllable. Then the control set D with nonvoid interior of system (9) is bounded if and only if A is hyperbolic.*

Proof By Theorem 15 there are compact sets $K \subset E^s, F \subset E^u$ such that

$$\overline{\mathbf{R}(0)} = K + E^c + E^u \text{ and } \overline{\mathbf{C}(0)} = F + E^c + E^s.$$

By Proposition 6, $D = \overline{\mathbf{R}(0)} \cap \mathbf{C}(0)$, because $0 \in D_0 \subset \text{int}D$, and hence every element $x \in D$ can be represented in the following two ways:

$$x = k + x_1 + x_+ = f + x_1 + x_-,$$

where $k \in K \subset E^s, f \in F \subset E^u, x_1 \in E^c, x_- \in E^s$ and $x_+ \in E^u$. Since $\mathbb{R}^d = E^s \oplus E^c \oplus E^u$ we get $k = x_-, f = x_+$. As $E^c = E^{sc} \cap E^{uc} \subset \mathbf{R}(0) \cap \mathbf{C}(0) \subset D$, we conclude that $E^c \subset D \subset K + E^c + F$, and so the control set D is bounded if and only if $E^c = \{0\}$. \square

Remark 20 We know that in the hyperbolic case

$$D = K_0 + F' \tag{10}$$

with $K_0 \subset E^s, F' \subset F \subset E^u$, where K_0 and F are compact sets with $0 \in K_0 \cap F$. In particular, it follows that $K_0, F' \subset D$.

Next we present a simple example illustrating control sets.

Example 21 Consider for $d = 2$ and $m = 1$

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_k, \quad u_k \in U = [-1, 1].$$

We claim that for this hyperbolic matrix A the unique control set with nonvoid interior is $D = (-1, 1) \times [-2, 2]$. The stable subspace associated with the eigenvalue $\frac{1}{2}$ of A

is the y -axis, the unstable subspace associated with the eigenvalue 2 is the x -axis. For a constant control $u \in [-1, 1]$, one computes the equilibrium as $(x(u), y(u))^T = (u, 2u)^T$. In particular, for $u = 1$ and $u = -1$ one obtains the equilibria

$$\begin{bmatrix} x(1) \\ y(1) \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \text{ and } \begin{bmatrix} x(-1) \\ y(-1) \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix},$$

resp. It is clear that for all $u \in (-1, 1)$ the equilibrium $(-u, 2u)^T$ is in the interior of the control set D . Furthermore, observe that for $x_0 > 1$ one has in the next step $2x_0 + u > x_0$ and for $x_0 < -1$ one has $2x_0 + u < x_0$. If $y_0 > 2$, then $\frac{1}{2}y_0 + u < \frac{1}{2}y_0 + 1 \leq y_0$ and if $y_0 < -2$, then $\frac{1}{2}y_0 + u \geq \frac{1}{2}y_0 - 1 > y_0$. Hence solutions starting left of the vertical line $x = -1$ and right of $x = 1$ have to go to the left and to the right, respectively. Solutions which start above the horizontal line $y = 2$ and below $y = -2$, have to go down and up, respectively. This shows that the control set must be contained in $(-1, 1) \times [-2, 2]$. The controllability property within D can be seen by the following analysis. If we start in an equilibrium $(x(\alpha), y(\alpha))^T = (-\alpha, 2\alpha)^T, \alpha \in (-1, 1)$, we get e.g.

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} -2\alpha \\ \alpha \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_0, \quad \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -4\alpha \\ \frac{1}{2}\alpha \end{bmatrix} + \begin{bmatrix} 2 \\ \frac{1}{2} \end{bmatrix} u_0 + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_1.$$

For the reachable set, we see that after one step the line segment $S = \{(u, u)^T, u \in [-1, 1]\}$ is shifted to $(-2\alpha, \alpha)^T$. After two time steps the line segment S is shifted to $(-4\alpha, \frac{1}{2}\alpha)^T$ and at every point the line segment $\{(2u, \frac{1}{2}u)^T | u \in [-1, 1]\}$ is added. One can show that the equilibrium $(0, 0)^T$ can be reached. If we start in $(0, 0)^T$, we compute

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_0, \quad \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 \\ \frac{1}{2} \end{bmatrix} u_0 + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_1, \\ \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} 4 \\ \frac{1}{4} \end{bmatrix} u_0 + \begin{bmatrix} 2 \\ \frac{1}{2} \end{bmatrix} u_1 + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_2.$$

Proceeding in this way one finds that one can get approximately to all points in D and, in particular, to the equilibria $(-1, 2)^T$ and $(1, -2)^T$. Connecting appropriately the controls, one finally shows that $D = (-1, 1) \times [-2, 2]$ is a control set.

5 Invariance Pressure

In this section we recall the concept of invariance pressure considered in [1,2,18] where potentials are defined on the control range. Furthermore, we introduce the generalized version of total invariance pressure, where the potentials are defined on the product of the state space and the control range. Again we consider the general system (1).

A pair (K, Q) of nonvoid subsets of M is called admissible if $K \subset Q$ is compact and for each $x \in K$ there exists $u \in \mathcal{U}$ such that $\varphi(\mathbb{N}, x, u) \subset Q$. For an admissible pair (K, Q) and $\tau > 0$, a (τ, K, Q) -spanning set \mathcal{S} of controls is a subset of \mathcal{U} such that for all $x \in K$ there is $u \in \mathcal{S}$ with $\varphi(k, x, u) \in Q$ for all $k \in \{1, \dots, \tau\}$. Denote by $C(U, \mathbb{R})$ the set of continuous function $f : U \rightarrow \mathbb{R}$ which we call potentials.

For a potential $f \in C(U, \mathbb{R})$ denote $(S_\tau f)(u) := \sum_{i=0}^{\tau-1} f(u_i), u \in \mathcal{U}$, and

$$a_\tau(f, K, Q) = \inf \left\{ \sum_{u \in \mathcal{S}} e^{(S_\tau f)(u)} \mid \mathcal{S} (\tau, K, Q)\text{-spanning} \right\}.$$

Definition 22 The invariance pressure $P_{inv}(f, K, Q)$ of control system (1) is defined by

$$P_{inv}(f, K, Q) = \overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \log a_{\tau}(f, K, Q).$$

For the potential $f = \mathbf{0}$, this reduces to the notion of invariance entropy, $P_{inv}(\mathbf{0}, K, Q) = h_{inv}(K, Q)$.

In order to define the total invariance pressure associate to every control u in a (τ, K, Q) -spanning set \mathcal{S} of controls an initial value $x_u \in K$ with $\varphi(k, x_u, u) \in Q$ for all $k \in \{1, \dots, \tau\}$. Then a set of state-control pairs of the form

$$\mathcal{S}_{tot} = \{(x_u, u) \in K \times \mathcal{S} \mid \varphi(k, x_u, u) \in Q \text{ for all } k \in \{1, \dots, \tau\}\}$$

is called totally (τ, K, Q) -spanning. Denote by $C(Q \times U, \mathbb{R})$ the set of continuous function $f : Q \times U \rightarrow \mathbb{R}$ which we again call potentials. For a potential $f \in C(Q \times U, \mathbb{R})$ and $(x, u) \in M \times \mathcal{U}$ denote $(S_{\tau} f)(x, u) := \sum_{i=0}^{\tau-1} f(\varphi(i, x, u), u_i)$ and

$$a_{\tau}(f, K, Q) := \inf \left\{ \sum_{(x,u) \in \mathcal{S}_{tot}} e^{(S_{\tau} f)(x,u)} \mid \mathcal{S}_{tot} \text{ totally } (\tau, K, Q)\text{-spanning} \right\}.$$

Definition 23 The total invariance pressure $P_{tot}(f, K, Q; \Sigma)$ of control system (1) is defined by

$$P_{tot}(f, K, Q) = \overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \log a_{\tau}(f, K, Q). \tag{11}$$

Note that by continuity and monotonicity of the logarithm,

$$P_{tot}(f, K, Q) = \overline{\lim}_{\tau \rightarrow \infty} \inf \left\{ \frac{1}{\tau} \log \sum_{(x,u) \in \mathcal{S}_{tot}} e^{(S_{\tau} f)(x,u)} \mid \mathcal{S}_{tot} \text{ totally } (\tau, K, Q)\text{-spanning} \right\}. \tag{12}$$

Furthermore $-\infty < a_{\tau}(f, K, Q) \leq \infty$ for every $\tau \in \mathbb{N}$, every admissible pair (K, Q) , and every potential f if every countable totally spanning set contains a finite totally spanning subset, cf. [2, Remark 7]. If $f(x, u)$ is independent of x , i.e., it is a continuous function on U , the total invariance pressure coincides with the invariance pressure.

Remark 24 The definition of totally (τ, K, Q) -spanning sets is inspired by the definition of spanning sets for (K, Q) in Wang, Huang, and Sun [15, p. 313], where a similar notion is introduced in the context of invariant partitions which provide an alternative definition of invariance entropy.

The next elementary proposition presents some properties of the function $P_{tot}(\cdot, K, Q) : C(Q \times U, \mathbb{R}) \rightarrow \mathbb{R} \cup \{\pm\infty\}$.

Proposition 25 *The following assertions hold for an admissible pair (K, Q) , functions $f, g \in C(Q \times U, \mathbb{R})$ and $c \in \mathbb{R}$:*

- (i) For $f \leq g$ one has $P_{tot}(f, K, Q) \leq P_{tot}(g, K, Q)$.
- (ii) $P_{tot}(f + c, K, Q) = P_{tot}(f, K, Q) + c$.

Proof This follows easily from the definition, cf. also [1, Proposition 13]. □

The following proposition shows that, in the definition of total invariance pressure, we can take the limit superior over times which are integer multiples of some fixed time step $\tau \in \mathbb{N}$. The proof is analogous to the proof given in [2, Theorem 20] for invariance pressure of continuous-time systems.

Proposition 26 *For all $f \in C(Q \times U, \mathbb{R})$ with $\inf_{(x,u) \in Q \times U} f(x, u) > -\infty$ the total invariance pressure satisfies for $\tau \in \mathbb{N}$*

$$P_{tot}(f, K, Q) = \overline{\lim}_{n \rightarrow \infty} \frac{1}{n\tau} \log a_{n\tau}(f, K, Q).$$

Proof For every $f \in C(Q \times U, \mathbb{R})$, the inequality

$$P_{tot}(f, K, Q) \geq \overline{\lim}_{n \rightarrow \infty} \frac{1}{n\tau} \log a_{n\tau}(f, K, Q) \tag{13}$$

is obvious. For the converse note that the function $g(x, u) := f(x, u) - \inf f$ is nonnegative (if $f \geq 0$, we may consider f instead of g). Let $\tau_k \in (0, \infty)$ with $\tau_k \rightarrow \infty$ for $k \rightarrow \infty$. Then for every $k \geq 1$ there exists $n_k \in \mathbb{N}_0$ such that $n_k \tau \leq \tau_k < (n_k + 1)\tau$ and $n_k \rightarrow \infty$ for $k \rightarrow \infty$. Since $g \geq 0$ it follows that

$$a_{\tau_k}(g, K, Q) \leq a_{(n_k+1)\tau}(g, K, Q)$$

and consequently

$$\frac{1}{\tau_k} \log a_{\tau_k}(g, K, Q) \leq \frac{1}{n_k \tau} \log a_{(n_k+1)\tau}(g, K, Q).$$

This yields

$$\overline{\lim}_{k \rightarrow \infty} \frac{1}{\tau_k} \log a_{\tau_k}(g, K, Q) \leq \overline{\lim}_{k \rightarrow \infty} \frac{1}{n_k \tau} \log a_{(n_k+1)\tau}(g, K, Q).$$

Since $\frac{1}{n_k \tau} = \frac{n_k+1}{n_k} \frac{1}{(n_k+1)\tau}$ and $\frac{n_k+1}{n_k} \rightarrow 1$ for $k \rightarrow \infty$, we obtain

$$\begin{aligned} \overline{\lim}_{k \rightarrow \infty} \frac{1}{\tau_k} \log a_{\tau_k}(g, K, Q) &\leq \overline{\lim}_{k \rightarrow \infty} \frac{1}{(n_k + 1)\tau} \log a_{(n_k+1)\tau}(g, K, Q) \\ &\leq \overline{\lim}_{n \rightarrow \infty} \frac{1}{n\tau} \log a_{n\tau}(g, K, Q). \end{aligned}$$

Together with Proposition 25 (ii) and (13) applied to $f - \inf f$, this shows that

$$\begin{aligned} P_{tot}(f, K, Q) &= P_{tot}(f - \inf f, K, Q) + \inf f \\ &= \overline{\lim}_{n \rightarrow \infty} \frac{1}{n\tau} \log a_{n\tau}(f - \inf f, K, Q) + \inf f \\ &= \overline{\lim}_{n \rightarrow \infty} \frac{1}{n\tau} \log a_{n\tau}(f, K, Q). \end{aligned}$$

□

The following result is given in [2, Corollary 15] for continuous-time systems. The discrete-time case is proved analogously.

Proposition 27 *Let K_1, K_2 be two compact sets with nonvoid interior contained in a control set $D \subset M$ and assume that every point in D is accessible. Then (K_1, D) and (K_2, D) are admissible pairs and for all $f \in C(U, \mathbb{R})$ we have*

$$P_{inv}(f, K_1, D) = P_{inv}(f, K_2, D).$$

6 Invariance Pressure for Linear Systems

The main result of this section presents a formula for the invariance pressure of the unique control set with nonvoid interior for hyperbolic linear control systems of the form (9).

We start with a proposition providing an upper bound for the total invariance pressure of the unique control set with nonvoid interior, cf. Theorems 18 and 19. The proof uses arguments from [3] which in turn are based on a construction by Kawan [9, Theorem 4.3], [10, Theorem 5.1] (for the discrete-time case cf. also [10, Remark 5.4] and Nair, Evans, Mareels, Moran [12, Theorem 3]).

Let A^+ be the restriction of A to the unstable subspace E^u . The unstable determinant of A is

$$\det A^+ = \prod_{\lambda \in \sigma(A)} \lambda^{n_\lambda} \text{ and } \log |\det A^+| = \sum_{\lambda \in \sigma(A)} n_\lambda \max\{0, \log |\lambda|\},$$

where n_λ denotes the algebraic multiplicity of an eigenvalue λ of A .

Proposition 28 *Consider a linear control system of the form (9) and assume that the pair (A, B) is controllable with a hyperbolic matrix A . Let D be the unique control set with nonvoid interior and let $f \in C(\overline{D} \times U, \mathbb{R})$. Then there exists a compact set $K \subset D$ with nonvoid interior such that the total invariance pressure satisfies*

$$P_{tot}(f, K, D) \leq \log |\det A^+| + \inf_{(\tau, x, u)} \frac{1}{\tau} \sum_{i=0}^{\tau-1} f(\varphi(i, x, u), u_i),$$

where the infimum is taken over all $\tau \in \mathbb{N}$ with $\tau \geq d$ and all τ -periodic controls u with a τ -periodic trajectory $\varphi(\cdot, x, u)$ in $\text{int} D$ such that $u_i \in \text{int} U$ for $i \in \{0, \dots, \tau - 1\}$.

Proof We will construct a compact subset $K \subset D$ with nonvoid interior such that the inequality above holds. Observe that then by Proposition 27 the pair (K, D) is admissible.

We may suppose that A has real Jordan form $R = T^{-1}AT$. In fact, writing $x = Tx'$ one obtains

$$x'_{k+1} = T^{-1}ATx'_k + T^{-1}Bu_k = Rx'_k + B'u_k \tag{14}$$

with $B' := T^{-1}B$. Then with $f'(x', u) = f(Tx', u) =: f(x, u)$, $K' := T^{-1}K$, and $D' := T^{-1}D$ the total invariance pressure $P_{tot}(f, K, D)$ coincides with the total invariance pressure $P_{tot}(f', K', D')$ of (14). Consider a τ^0 -periodic control $u^0(\cdot)$ with τ^0 -periodic trajectory $\varphi(\cdot, x^0, u^0)$ as in the statement of the theorem, hence

$$x^0 = R^{\tau^0}x^0 + \sum_{i=0}^{\tau^0-1} R^{\tau^0-i}B'u_i. \tag{15}$$

Step 1: Choose a basis \mathcal{B} of \mathbb{R}^d adapted to the real Jordan structure of R and let $L_1(R), \dots, L_r(R)$ be the Lyapunov spaces of R , that is, the sums of the generalized eigenspaces corresponding to eigenvalues λ with the absolute value $|\lambda| = \rho_j$. This yields the decomposition

$$\mathbb{R}^d = L_1(R) \oplus \dots \oplus L_r(R).$$

Let $d_j = \dim L_j(R)$ and denote the restriction of R to $L_j(R)$ by R_j . Now take an inner product on \mathbb{R}^d such that the basis \mathcal{B} is orthonormal with respect to this inner product and let $\|\cdot\|$ denote the induced norm.

Step 2: We fix some constants: Let S_0 be a real number which satisfies

$$S_0 > \sum_{j=1}^r \max\{1, d_j \rho_j\} = \log |\det A^+|,$$

and choose $\xi = \xi(S_0) > 0$ such that

$$0 < d\xi < S_0 - \sum_{j=1}^r \max\{1, d_j \rho_j\}$$

and such that $\rho_j < 1$ implies $\rho_j + \xi < 1$ for all j . Let $\delta \in (0, \xi)$. It follows that there exists a constant $c = c(\delta) \geq 1$ such that for all j and for all $k \in \mathbb{N}$

$$\|R_j^k\| \leq c(\rho_j + \delta)^k.$$

For every $m \in \mathbb{N}$ we define positive integers by

$$M_j(m) := \begin{cases} \lfloor (\rho_j + \xi)^m \rfloor + 1 & \text{if } \rho_j \geq 1 \\ 1 & \text{if } \rho_j < 1 \end{cases}$$

and a function $\beta : \mathbb{N} \rightarrow (0, \infty)$ by

$$\beta(m) := \max_{1 \leq j \leq r} \left\{ (\rho_j + \delta)^m \frac{\sqrt{d_j}}{M_j(m)} \right\}, \quad m \in \mathbb{N}.$$

If $\rho_j < 1$, then $\rho_j + \delta < 1$ and $M_j(m) \equiv 1$, and hence $(\rho_j + \delta)^m / M_j(m)$ converges to zero for $m \rightarrow \infty$. If $\rho_j \geq 1$, we have $M_j(m) \geq (\rho_j + \xi)^m$ and hence

$$(\rho_j + \delta)^m \frac{\sqrt{d_j}}{M_j(m)} \leq (\rho_j + \delta)^m \frac{\sqrt{d_j}}{(\rho_j + \xi)^m} = \left(\frac{\rho_j + \delta}{\rho_j + \xi} \right)^m \sqrt{d_j}. \tag{16}$$

Since $\delta \in (0, \xi)$, we have $\frac{\rho_j + \delta}{\rho_j + \xi} < 1$ showing that also in this case $\beta(m) \rightarrow 0$ for $m \rightarrow \infty$.

Since we assume controllability of (A, B) and $\tau^0 \geq d$ there exists $C_0 > 0$ such that for every $x \in \mathbb{R}^d$ there is a control $u \in \mathcal{U}$ with

$$\varphi(\tau^0, x, u) = R^{\tau^0} x + \sum_{i=0}^{\tau^0-1} R^{\tau^0-i} B' u_i = 0 \text{ and } \|u\|_\infty \leq C_0 \|x\|. \tag{17}$$

The inequality follows by the inverse mapping theorem. For the corresponding trajectory we find a constant $C_1 > 0$ such that for $k \in \{1, \dots, \tau^0\}$

$$\|\varphi(k, x, u)\| \leq \|R\|^k \|x\| + \sum_{i=0}^{k-1} \|R\|^{k-i} \|B'\| C_0 \|x\| \leq C_1 \|x\|. \tag{18}$$

For $b_0 > 0$ let C be the d -dimensional compact cube C in \mathbb{R}^d centered at the origin with sides of length $2b_0$ parallel to the vectors of the basis B . Choose b_0 small enough such that

$$K := x^0 + C \subset D$$

and $\overline{B(u^0(k), Cb_0)} \subset U$ for all $k \in \{0, \dots, \tau^0\}$. This is possible, since $x^0 \in \text{int}D$ and all values $u^0(k)$ are in the interior of U .

Step 3. Let $\varepsilon > 0$ and $\tau = m\tau^0$ with $m \in \mathbb{N}$. By Theorem 19, the closure \overline{D} is compact, hence for the continuous function f on the compact set $\overline{D} \times U$ there is $\varepsilon_1 > 0$ such that for all $(x, u), (x', u') \in \overline{D} \times U$

$$\max \{ \|x - x'\|, \|u - u'\| \} < \varepsilon_1 \text{ implies } |f(x, u) - f(x', u')| < \varepsilon. \tag{19}$$

We may take $m \in \mathbb{N}$ large enough such that

$$\frac{d}{\tau} \log 2 = \frac{d}{m\tau^0} \log 2 < \varepsilon. \tag{20}$$

Furthermore, we may choose b_0 small enough such that

$$C_0 b_0 < \varepsilon_1 \text{ and } C_1 b_0 < \varepsilon_1. \tag{21}$$

Partition \mathcal{C} by dividing each coordinate axis corresponding to a component of the j th Lyapunov space $L_j(R)$ into $M_j(\tau)$ intervals of equal length. The total number of subcuboids in this partition of \mathcal{C} is $\prod_{j=1}^r M_j(\tau)^{d_j}$. Next we will show that it suffices to take $\prod_{j=1}^r M_j(\tau)^{d_j}$ control functions to steer the system from all states in $x^0 + \mathcal{C}$ back to $x^0 + \mathcal{C}$ in time τ such that the controls are within distance ε_1 to u^0 and the corresponding trajectories remain within distance ε_1 from the trajectory $\varphi(\cdot, x^0, u^0)$. Let y be the center of a subcuboid. By (17) there exists $u = (u_0, \dots, u_{\tau^0-1})$ such that

$$\varphi(\tau^0, y, u) = 0 \text{ and } \|u\|_\infty \leq C_0 \|y\| \leq C_0 b_0 < \varepsilon_1. \tag{22}$$

For $k \geq t_0$ let $u_k = 0$. Hence $\varphi(\tau, y, u) = 0$ and $u(t) \in U$ for all $k \in \{0, \dots, \tau\}$. Using (15) and linearity, we find that $x^0 + y$ is steered by $u^0 + u$ in time $\tau = m\tau^0$ to x^0 ,

$$\varphi(\tau, x^0 + y, u^0 + u) = \varphi(\tau, x^0, u^0) + \varphi(\tau, y, u) = x^0. \tag{23}$$

Now consider an arbitrary point $x \in \mathcal{C}$. Then it lies in one of the subcuboids and we denote the corresponding center of this subcuboid by y with associated control $u = u(y)$. We will show in **Step 4** that $u^0 + u$ also steers $x^0 + x$ back to $x^0 + \mathcal{C}$ and in **Step 5** that the corresponding trajectory $\varphi(k, x^0 + x, u^0 + u)$ remains within distance ε_1 of $\varphi(k, x^0, u^0)$, $k \in \{0, \dots, \tau\}$.

Step 4. Observe that

$$\|x - y\| \leq \frac{b_0}{M_j(\tau)} \sqrt{d_j}.$$

By (16) this implies that

$$\|R^\tau x - R^\tau y\| \leq \|R_j^{m\tau^0}\| \|x - y\| \leq c(\rho_j + \delta)^{m\tau^0} \frac{b_0}{M_j(m\tau^0)} \sqrt{d_j} \rightarrow 0 \text{ for } m \rightarrow \infty,$$

and hence for m large enough $\|R^\tau x - R^\tau y\| \leq b_0$. This implies that the solution $\varphi(k, x^0 + x, u^0 + u)$, $k \in \mathbb{N}$, satisfies for m large enough by (23) and linearity,

$$\begin{aligned} & \|\varphi(\tau, x^0 + x, u^0 + u) - x^0\| \\ &= \left\| R^\tau(x^0 + x) + \sum_{i=0}^{\tau-1} R^{\tau-i} B'(u_i^0 + u_i) - x^0 \right\| \\ &\leq \|R^\tau(x^0 + x) - R^\tau(x^0 + y)\| + \left\| R^\tau(x^0 + y) + \sum_{i=0}^{\tau-1} R^{\tau-i} B'(u_i^0 + u_i) - x^0 \right\| \\ &\leq \|R^\tau x - R^\tau y\| + \|\varphi(\tau, x^0 + y, u^0 + u) - x^0\| \end{aligned}$$

$$\leq b_0 + 0.$$

This shows that $\varphi(\tau, x^0 + x, u^0 + u) \in x^0 + C$ and it also follows that $\varphi(\tau, x^0 + x, u^0 + u) \in D$ for all $k \in \{0, 1, \dots, \tau\}$.

Step 5. By linearity and formulas (17), (18), and (21) we can estimate for $k \in \{0, 1, \dots, \tau^0\}$

$$\begin{aligned} & \|\varphi(k, x^0 + x, u^0 + u) - \varphi(k, x^0, u^0)\| \\ &= \left\| R^k(x^0 + x) + \varphi(k, 0, u^0 + u) - R^k x^0 - \varphi(k, 0, u^0) \right\| \\ &= \left\| R^k x + \varphi(k, 0, u) \right\| = \|\varphi(k, x, u)\| \leq C_1 \|x\| \leq C_1 b_0 < \varepsilon_1. \end{aligned}$$

Together with (22) and (19) this shows that for $k \in \{0, 1, \dots, \tau\}$

$$|f(\varphi(k, x^0 + x, u^0 + u), u_k^0 + u_k) - f(\varphi(k, x^0, u^0), u_k^0)| < \varepsilon. \tag{24}$$

Step 6. We have constructed $\prod_{j=1}^r M_j(\tau)^{d_j}$ control functions that allow us to steer the system from all states in $K = x^0 + C$ back to $x^0 + C$ in time τ and satisfy (24). By iterated concatenation of these control functions we obtain a totally $(n\tau, K, D)$ -spanning set \mathcal{S}_{tot} for each $n \in \mathbb{N}$ with cardinality

$$\#\mathcal{S}_{tot} = \left(\prod_{j=1}^r M_j(\tau)^{d_j} \right)^n = \left(\prod_{j:\rho_j \geq 0} (\lfloor (\rho_j + \xi)^\tau \rfloor + 1)^{d_j} \right)^n.$$

By (24) it follows that

$$\begin{aligned} \log a_{n\tau}(f, K, D) &\leq \log \left(\sum_{(x,u) \in \mathcal{S}_{tot}} e^{(S_{n\tau} f)(x,u)} \right) \\ &= \log \left(\sum_{(x,u) \in \mathcal{S}_{tot}} e^{(S_{n\tau} f)(x^0, u^0)} \cdot e^{(S_{n\tau} f)(x,u) - (S_{n\tau} f)(x^0, u^0)} \right) \\ &\leq \log \sum_{(x,u) \in \mathcal{S}_{tot}} e^{(S_{n\tau} f)(x^0, u^0)} + \log e^{\sum_{i=0}^{n\tau-1} \varepsilon} \\ &\leq \log \left(\#\mathcal{S}_{tot} \cdot e^{(S_{n\tau} f)(x^0, u^0)} \right) + n\tau\varepsilon. \end{aligned}$$

This implies, using also (20),

$$\begin{aligned} \frac{1}{n\tau} \log a_{n\tau}(f, K, D) &\leq \frac{1}{\tau} \sum_{j:\rho_j \geq 0} d_j \log(\lfloor e^{(\rho_j + \xi)^\tau} \rfloor + 1) + \frac{1}{n\tau} \sum_{i=0}^{n\tau-1} f(\varphi(i, x^0, u^0), u_i^0) + \varepsilon \\ &\leq \frac{1}{\tau} \sum_{j:\rho_j \geq 0} d_j \log(2e^{(\rho_j + \xi)^\tau}) + \frac{1}{\tau^0} \sum_{i=0}^{\tau^0-1} f(\varphi(i, x^0, u^0), u_i^0) + \varepsilon \\ &\leq \frac{d}{\tau} \log 2 + \frac{1}{\tau} \sum_{j:\rho_j \geq 0} d_j (\rho_j + \xi)\tau + \frac{1}{\tau^0} \sum_{i=0}^{\tau^0-1} f(\varphi(i, x^0, u^0), u_i^0) + \varepsilon \\ &\leq \varepsilon + d\xi + \sum_{j:\rho_j \geq 0} d_j \rho_j + \frac{1}{\tau^0} \sum_{i=0}^{\tau^0-1} f(\varphi(i, x^0, u^0), u_i^0) + \varepsilon \\ &< S_0 + \frac{1}{\tau^0} \sum_{i=0}^{\tau^0-1} f(\varphi(i, x^0, u^0), u_i^0) + 2\varepsilon. \end{aligned}$$

Since ε can be chosen arbitrarily small and S_0 arbitrarily close to $\log |\det A^+|$, the assertion of the proposition follows. □

For the invariance pressure, we obtain the following consequence.

Corollary 29 *Consider a linear control system of the form (9) and assume that the pair (A, B) is controllable with a hyperbolic matrix A . Let D be the unique control set with nonvoid interior and let $f \in C(U, \mathbb{R})$. Then for every compact set $K \subset D$ with nonvoid interior the invariance pressure satisfies*

$$P_{inv}(f, K, D) \leq \log |\det A^+| + \inf_{(\tau, x, u)} \frac{1}{\tau} \sum_{i=0}^{\tau-1} f(u_i),$$

where the infimum is taken over all $\tau \in \mathbb{N}$ with $\tau \geq d$ and all τ -periodic controls u with a τ -periodic trajectory $\varphi(\cdot, x, u)$ in $\text{int}D$ such that $u_i \in \text{int}U$ for $i \in \{0, \dots, \tau - 1\}$.

Proof The assertion follows from Proposition 28, since every compact subset of D is contained in a compact subset K of D with nonvoid interior and the invariance pressure is independent of the choice of such a set K by Proposition 27. □

Remark 30 Kawan [10, Theorem 3.1] derives for the outer invariance entropy $h_{inv,out}(K, Q)$, which is a lower bound for the invariance entropy, the formula

$$h_{inv,out}(K, Q) = \log |\det A^+|.$$

Here (K, Q) is an admissible pair, K has positive Lebesgue measure, and Q is compact. For the potential $f = 0$, Corollary 29 shows that the invariance entropy satisfies

$$h_{inv}(K, Q) \leq \log |\det A^+| = h_{inv,out}(K, Q) \leq h_{inv}(K, Q)$$

implying that

$$h_{inv}(K, Q) = \log |\det A^+|. \tag{25}$$

We proceed to prove a lower bound for the invariance pressure. Recall that with respect to A the state space \mathbb{R}^d can be decomposed into the direct sum of the center-stable subspace E^{sc} and the unstable subspace E^u which are the direct sums of all generalized real eigenspaces for the eigenvalues λ with $|\lambda| \leq 1$ and $|\lambda| > 1$, resp. Let $\pi : \mathbb{R}^d \rightarrow E^u$ be the projection along E^{sc} .

Proposition 31 *Let $K \subset D$ be compact and assume that both K and D have positive and finite Lebesgue measure. Then for every $f \in C(U, \mathbb{R})$*

$$P_{inv}(f, K, D) \geq \log |\det A^+| + \inf_{(\tau, x, u)} \frac{1}{\tau} \sum_{i=0}^{\tau-1} f(u_i),$$

where the infimum is taken over all $(\tau, x, u) \in \mathbb{N} \times D \times U$ with $\tau \geq d$ and $\pi\varphi(i, x, u) \in \pi D$ for $i \in \{0, 1, \dots, \tau - 1\}$.

Proof Every (τ, K, Q) -spanning set \mathcal{S} satisfies

$$\log \sum_{u \in \mathcal{S}} e^{(S_\tau f)(u)} \geq \log \inf_{u \in \mathcal{S}} e^{(S_\tau f)(u)} + \log \#\mathcal{S}. \tag{26}$$

First suppose that the unstable subspace of A is trivial, $E^u = 0$. Formula (25) implies that

$$\overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \inf \{ \log \# \mathcal{S} \mid \mathcal{S}(\tau, K, Q)\text{-spanning} \} = h_{inv}(K, D) = \log |\det A^+| = 0.$$

Now (12) and (26) implies

$$\begin{aligned} P_{inv}(f, K, Q) &= \overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \inf \left\{ \log \sum_{u \in \mathcal{S}} e^{(S_\tau f)(u)} \mid \mathcal{S}(\tau, K, Q)\text{-spanning} \right\} \\ &\geq \overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \inf \left\{ \log \inf_{u \in \mathcal{S}} e^{(S_\tau f)(u)} + \log \# \mathcal{S} \mid \mathcal{S}(\tau, K, Q)\text{-spanning} \right\} \\ &\geq \overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \inf \left\{ \inf_{u \in \mathcal{S}} \sum_{i=0}^{\tau-1} f(u_i) \mid \mathcal{S}(\tau, K, Q)\text{-spanning} \right\} + 0 \\ &\geq \overline{\lim}_{\tau \rightarrow \infty} \inf_{u \in \mathcal{S}} \frac{1}{\tau} \sum_{i=0}^{\tau-1} f(u_i) \geq \inf_{u \in \mathcal{S}} \frac{1}{\tau} \sum_{i=0}^{\tau-1} f(u_i). \end{aligned}$$

Since for $u \in \mathcal{S}$ there is $x \in K$ with $\pi\varphi(i, x, u) = 0 \in \pi D$ for $i \in \{0, 1, \dots, \tau - 1\}$, the assertion for trivial unstable subspace E^- follows.

Now suppose that E^u is nontrivial. We may assume that $P_{inv}(f, K, Q) < \infty$ and hence and all considered spanning sets are countable. Note that by invariance of E^{sc} and E^u the induced system on E^u is well defined with trajectories $\pi\varphi(k, x, u)$, $k \in \mathbb{N}$. For each u in a (τ, K, D) -spanning set \mathcal{S} define

$$\pi K_u := \pi K \cap \bigcap_{i=0}^{\tau-1} (\pi\varphi_{i,u})^{-1}(D).$$

Thus $\pi K = \bigcup_{u \in \mathcal{S}} \pi K_u$. Since D is measurable, each set πK_u is measurable as the countable intersection of measurable sets. We denote the Lebesgue measure in \mathbb{R}^d by μ^d and the induced measure on E^u by μ . The linear part of the affine-linear map $\pi\varphi_{\tau,u}(x)$ is given by $(A^+)^{\tau}$, hence it follows that

$$\mu(\pi D) \geq \mu(\pi\varphi_{\tau,u}(\pi K_u)) = \int_{\pi\varphi_{\tau,u}(\pi K_u)} d\mu = \int_{\pi K_u} |\det(A^+)^{\tau}| d\mu = \mu(\pi K_u) |\det A^+|^{\tau}.$$

Abbreviate $\beta(\tau) = \inf_{(x,u)} (S_\tau f)(u)$, where the infimum is taken over all $(\pi x, u) \in \pi K \times \mathcal{U}$ with $\pi\varphi(i, x, u) \in \pi D$ for $i = 0, \dots, \tau - 1$. Then we find

$$\begin{aligned} e^{\beta(\tau)} \mu(\pi K) &\leq \sum_{u \in \mathcal{S}} e^{(S_\tau f)(u)} \mu(\pi K_u) \leq \sup_{u \in \mathcal{S}} \mu(\pi K_u) \sum_{u \in \mathcal{S}} e^{(S_\tau f)(u)} \\ &\leq \frac{\mu(\pi D)}{|\det A^+|^{\tau}} \sum_{u \in \mathcal{S}} e^{(S_\tau f)(u)}. \end{aligned}$$

Since this holds for every (τ, K, D) -spanning set \mathcal{S} and $\mu^d(D) > 0$ implies $\mu(\pi D) > 0$, we find

$$a_\tau(f, K, D) = \inf_{u \in \mathcal{S}} \{ \sum_{i=0}^{\tau-1} e^{(S_i f)(u)} \mid \mathcal{S}(\tau, K, D)\text{-spanning} \} \geq \frac{\mu(\pi K)}{\mu(\pi D)} e^{\beta(\tau)} |\det A^+|^{\tau},$$

implying

$$\begin{aligned}
 P_{inv}(f, K, D) &= \overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \log a_\tau(f, K, D) \geq \inf_\tau \frac{1}{\tau} \beta(\tau) + \log |\det A^+| \\
 &= \inf_{(\tau, x, u)} \frac{1}{\tau} (S_\tau f)(u) + \log |\det A^+|,
 \end{aligned}$$

where the infimum is taken over all $(\tau, x, u) \in \mathbb{N} \times \pi K \times \mathcal{U}$ with $\pi\varphi(i, x, u) \in \pi D$ for $i = 0, \dots, \tau - 1$. □

The next theorem is the main result of this paper. For linear discrete-time control systems it provides a formula for the invariance pressure of control sets.

Theorem 32 *Consider a linear control system of the form (9) and assume that the system without control restriction is controllable in \mathbb{R}^d , the matrix A is hyperbolic, and the control range U is a compact convex neighborhood of the origin with $U = \text{int}U$. Let D be the unique control set with nonvoid interior. Then D is bounded and for every compact set $K \subset D$ with nonvoid interior and every potential $f \in C(U, \mathbb{R})$, the invariance pressure is given by*

$$P_{inv}(f, K, D) = \log |\det A^+| + \min_{u \in U} f(u) = h_{inv}(K, D) + \min_{u \in U} f(u).$$

Proof Theorems 18 and 19 imply existence, uniqueness, and boundedness of the control set D . Formula (25) implies that $h_{inv}(K, D) = \log \det A^+$ showing the second equality above. Proposition 31 and Corollary 29 yield the bounds,

$$\inf_{(\tau', x', u')} \frac{1}{\tau'} \sum_{i=0}^{\tau'-1} f(u'_i) \leq P_{inv}(f, K, Q) - \log |\det A^+| \leq \inf_{(\tau, x, u)} \frac{1}{\tau} \sum_{i=0}^{\tau-1} f(u_i), \quad (27)$$

where the first infimum is taken over all $(\tau', x', u') \in \mathbb{N} \times D \times \mathcal{U}$ with $\tau' \geq d$ and $\pi\varphi(i, x', u') \in \pi D$ for $i \in \{0, \dots, \tau' - 1\}$ and the second infimum is taken over all $\tau \in \mathbb{N}$ with $\tau \geq d$ and all τ -periodic controls u with a τ -periodic trajectory $\varphi(\cdot, x, u)$ in $\text{int}D$ such that $u_i \in \text{int}U$ for $i \in \{0, \dots, \tau - 1\}$.

Note that there is a control value $u^0 \in U$ with $f(u^0) = \min_{u \in U} f(u)$. Consider

$$f(u^0) = \frac{1}{d} \sum_{i=0}^{d-1} f(u^0) \leq \inf_{(\tau', x', u')} \frac{1}{\tau'} \sum_{i=0}^{\tau'-1} f(u'_i), \quad (28)$$

where the infimum is taken over all triples $(\tau', x', u') \in \mathbb{N} \times K \times \mathcal{U}$ with $\tau' \geq d$ and $\pi\varphi(i, x', u') \in \pi D$ for $i \in \{0, \dots, \tau' - 1\}$. Let $\varepsilon > 0$. Then there is a control function u^1 with values in a compact subset of $\text{int}U$ such that

$$\frac{1}{d} \sum_{i=0}^{d-1} f(u^1_i) \leq \frac{1}{d} \sum_{i=0}^{d-1} f(u^0) + \varepsilon. \quad (29)$$

By hyperbolicity of A the matrix $I - A^d$ is invertible, and hence there exists a unique solution x^1 of

$$(I - A^d)x^1 = \varphi(d, 0, u^1).$$

Now by linearity

$$x^1 = A^d x^1 + \varphi(d, 0, u^1) = \varphi(d, x^1, u^1).$$

Since the values of u^1 are in $\text{int}U$ and (A, B) is controllable, it follows that a neighborhood of x^1 can be reached in time d from x^1 . Analogously, x^1 can be reached from every point in a neighborhood of x^1 in time d . Hence in the intersection of these two neighborhoods every point can be steered in time $2d$ into every other point. This shows that x^1 is in the interior of the control set D , and the corresponding trajectory $\varphi(i, x^1, u^1), i \in \{0, \dots, d - 1\}$, remains by Proposition 7 in the interior of D . Extending u^1 to a d -periodic control again denoted by u^1 we find that the control-trajectory pair $(u^1(\cdot), \varphi(\cdot, x^1, u^1))$ is d -periodic, the trajectory is contained in $\text{int}D$ and all values u_i^1 are in a compact subset of $\text{int}U$. It follows that

$$\begin{aligned} \inf_{(\tau', x', u')} \frac{1}{\tau'} \sum_{i=0}^{\tau'-1} f(u'_i) &\stackrel{(28)}{\geq} f(u^0) = \frac{1}{d} \sum_{i=0}^{d-1} f(u^0) \stackrel{(29)}{\geq} \frac{1}{d} \sum_{i=0}^{d-1} f(u_i^1) - \varepsilon \\ &\geq \inf_{(\tau, x, u)} \frac{1}{\tau} \sum_{i=0}^{\tau-1} f(u_i) - \varepsilon, \end{aligned}$$

where the first infimum is taken over all triples $(\tau', x', u') \in \mathbb{N} \times K \times \mathcal{U}$ with $\tau' \geq d$ and $\pi\varphi(i, x', u') \in \pi D$ for $i \in \{0, \dots, \tau' - 1\}$ and the second infimum is taken over all $(\tau, x, u) \in \mathbb{N} \times D \times \mathcal{U}$ such that the control-trajectory pair $(u, \varphi(\cdot, x, u))$ is τ -periodic with $\tau \geq d$, the trajectory is contained in $\text{int}D$, and the control values u_i are in a compact subset of $\text{int}U$.

Using this in (27) we get

$$\inf_{(\tau', x', u')} \frac{1}{\tau'} \sum_{i=0}^{\tau'-1} f(u'_i) \leq P_{inv}(f, K, Q) - \log |\det A^+| \leq \inf_{(\tau', x', u')} \frac{1}{\tau'} \sum_{i=0}^{\tau'-1} f(u'_i) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the assertion of the theorem follows. □

Remark 33 For partially hyperbolic control systems, Da Silva and Kawan prove in [11] relations between invariance entropy and topological pressure for the unstable determinant. In contrast to our framework, they consider the topological pressure (with respect to the fibers) of associated random dynamical systems obtained by endowing the space of controls with shift invariant probability measures.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Colonius, F., Cossich, J.A.N., Santana, A.: Invariance pressure for control systems. *J. Dyn. Differ. Equ.* **31**(1), 1–23 (2019)
2. Colonius, F., Santana, A., Cossich, J.A.N.: Invariance pressure of control sets. *SIAM J. Control Optim.* **56**(6), 4130–4147 (2018)

3. Colonius, F., Cossich, J.A.N., Santana, A.: Bounds for invariance pressure. *J. Differ. Equ.* **268**, 7877–7896 (2020)
4. Colonius, F.: Invariance entropy, quasi-stationary measures and control sets. *Discrete Contin. Dyn. Syst.: DCDS-A* **38**(4), 2093–2123 (2018)
5. Da Silva, A., Kawan, C.: Invariance entropy of hyperbolic control sets. *Discrete Contin. Dyn. Syst.: DCDS-A* **36**(1), 97–136 (2016)
6. Da Silva, A., Kawan, C.: Lyapunov exponents and partial hyperbolicity of chain control sets on flag manifolds. *Isr. J. Math.* **232**, 947–1000 (2019)
7. Hinrichsen, D., Pritchard, A.J.: *Mathematical Systems Theory*, vol. 2. Springer, Berlin (2021). (**in preparation**)
8. Huang, Y., Zhong, X.: Carathéodory–Pesin structures associated with control systems. *Syst. Control Lett.* **112**, 36–41 (2018)
9. Kawan, C.: Invariance entropy of control sets. *SIAM J. Control Optim.* **49**, 732–751 (2011)
10. Kawan, C.: *Invariance Entropy for Deterministic Control Systems. An Introduction*. LNM, vol. 2089. Springer, Berlin (2013)
11. Kawan, C., Da Silva, A.: Invariance entropy for a class of partially hyperbolic sets. *Math. Control Signals Syst.* (2018). <https://doi.org/10.1007/s00498-018-0224-2>
12. Nair, G., Evans, R.J., Mareels, I., Moran, W.: Topological feedback entropy and nonlinear stabilization. *IEEE Trans. Autom. Control* **49**, 1585–1597 (2004)
13. Patrão, M., San-Martin, L.: Semiflows on topological spaces: chain transitivity and semigroups. *J. Dyn. Differ. Equ.* **19**, 155–180 (2007)
14. Sontag, E.: *Mathematical Control Theory. Deterministic Finite Dimensional Systems*, 2nd edn. Springer, New York (1998)
15. Wang, T., Huang, Y., Sun, H.-W.: Measure-theoretic invariance entropy for control systems. *SIAM J. Control Optim.* **57**(1), 310–333 (2019)
16. Wing, J., Desoer, C.A.: The multiple-input minimal-time regulator problem (general theory). *IEEE Trans. Autom. Control: AC* **8**(2), 125–136 (1963)
17. Wirth, F.: Dynamics and controllability of nonlinear discrete-time control systems. *IFAC Proc. Vol.* **31**, 267–272 (1998)
18. Zhong, X., Huang, Y.: Invariance pressure dimensions for control systems. *J. Dyn. Differ. Equ.* **31**, 2205–2222 (2019)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.