Universität Augsburg

Institut für Mathematik

Fritz Colonius, Uwe Helmke

Entropy of Controlled-Invariant Subspaces

Preprint Nr. 31/2009 — 13. November 2009
Institut für Mathematik, Universitätsstraße, D-86135 Augsburg http://www.math.uni-augsburg.de/

Impressum:

Herausgeber:
Institut für Mathematik
Universität Augsburg
86135 Augsburg
`http://www.math.uni-augsburg.de/pages/de/forschung/preprints.shtml`

ViSdP:
Fritz Colonius
Institut für Mathematik
Universität Augsburg
86135 Augsburg

ENTROPY OF CONTROLLED–INVARIANT SUBSPACES*

FRITZ COLONIUS‡ AND UWE HELMKE¶

Abstract. For continuous-time linear control systems, a concept of entropy for controlled and almost controlled invariant subspaces is introduced. Upper bounds for the entropy in terms of the eigenvalues of the autonomous subsystem are derived.

September 15, 2009

Key words. invariance entropy, topological entropy, geometric control, almost (A,B)-invariant subspaces, eigenvalue bounds.

AMS subject classifications. 94A17, 37B40, 93C15

1. Introduction. Controlled and conditioned invariant subspaces of linear dynamical systems play a crucial role in understanding controller design problems such as disturbance decoupling, filtering, robust observer design, and high gain state feedback. In fact, starting form the early work of Basile–Morro [1] and Wonham [14], controlled invariant subspaces became a cornerstone of geometric control theory. In this paper, we begin an investigation of how geometric control design via controlled invariant subspaces is affected by entropy estimates and associated data rate constraints. The main motivation for this circle of ideas comes from the increasing needs of controlling systems with communication constraints, i.e. for systems where the state passes through a communication channel and may thus not be fully available to the controller.

As a starting point for such an investigation, we associate to any almost (A, B)–invariant subspace V of a linear control system a number, called the invariance entropy of V, that measures how difficult it is, using open loop controls, to keep the system in V. It is defined by the exponential growth rate of the number of controls necessary to keep the system in an arbitrarily small ε-neighborhood of V. More generally, by extending the familiar notion of topological entropy for the flow defined by A, we define the entropy of an arbitary linear subspace V of the state space. We show that the invariance entropy is finite for any almost (A, B)-invariant subspace and derive upper bounds in terms of the sum of the eigenvalues of A with positive real part. Sharper upper bounds are derived for specific classes of linear systems.

Our approach partially extends and follows that by Colonius and Kawan [4], where an entropy-like notion was proposed for controlled invariance of compact subsets of the state space of general control systems. Their approach in turn has been motivated by the work of Nair et al. [8] on feedback entropies for nonlinear discrete-time systems. The entropy notion considered here may be regarded as a lower bound for the minimum data rate (take the logarithm with base 2 instead of the natural logarithm used, for convenience, in the present paper.) More explicit relations to data rates are given in Kawan's PhD thesis [9].

The contents of this paper are as follows: Section 2 recalls basic facts on (A, B)-invariant and on almost (A, B)-invariant subspaces. Section 3 introduces entropy of

*Supported by DFG grants Co 124/17-1 and HE 1858/12-1 within DFG Priority Program 1305.

‡Institut für Mathematik, Universität Augsburg, Universitätsstraße, 86135 Augsburg/Germany, fritz.colonius@math.uni-augsburg.de

¶Fakultät für Mathematik, Universität Würzburg, Am Hubland, 97074 Würzburg/Germany, helmke@mathematik.uni-wuerzburg.de

almost (A,B)-invariant subspaces and provides an upper bound in terms of eigenvalues. We also show that this invariance entropy can be characterized by an entropy property of the uncontrolled system and use this to show sharper bounds for special cases. Since these results rely on Bowen's classical characterization [2] of entropy of linear maps, we have, for the convenience of the reader, included an essentially self-contained proof of his result in the appendix (where also related notation is recalled).

2. Preliminaries on controlled-invariant subspaces. The purpose of this section is to summarize some well-known definitions and facts from geometric control theory, i.e. controlled and almost controlled invariant subspaces. The notion of controlled invariant or (A,B)–invariant subspaces was introduced by Basile and Morro [1], and Wonham [14], while almost controlled invariant subspaces were first introduced by J.C. Willems [13]; we also refer to the PhD Thesis by J. Trumpf [12] for a useful summary of basic definitions and facts.

Consider linear control systems in state space form

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2.1}$$

with matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. The solutions of (2.1) are given by the variations-of-constants formula:

$$\varphi(t,x,u) = e^{At}x + \int_0^t e^{A(t-s)}Bu(s)ds.$$

Recall that a subspace V is called (A,B)-invariant, if for all $x \in V$ there is $u \in \mathbb{R}^m$ with

$$Ax + Bu \in V.$$

Equivalently, there is a matrix $F \in \mathbb{R}^{m\times n}$, a so-called friend of V, such that for $A_F := A + BF$

$$A_F V \subset V.$$

This can be seen by choosing for a basis $x_1, ..., x_k$ of V control values $u_1, ..., u_k \in \mathbb{R}^m$ with $Ax_i + Bu_i \in V$. Then extend this to a basis of $\mathbb{R}^n$ and define a linear map F by

$$Fx_i = u_i, \text{ for } i = 1, ..., k\,, \text{ and } F \text{ arbitrary outside } V.$$

This also shows that V is (A,B)-invariant if and only if it is controlled invariant, i.e., for every $x \in V$ there is an open loop continuous control function $u : \mathbb{R} \to \mathbb{R}^m$ with $\varphi(t,x,u) \in V$ for all $t \geq 0$. In fact, differentiating the solution one finds

$$V \ni \frac{d}{dt}\varphi(0,x,u) = Ax + Bu(0).$$

Conversely, define for $x \in V$ a control by $u(t) = Fe^{(A+BF)t}x, t \geq 0$.

A linear subspace V is called almost (A,B)-invariant, if for any $x \in V$ and any $\varepsilon > 0$ there exists a control function $u(\cdot)$ such that for all $t \geq 0$

$$\operatorname{dist}(\varphi(t,x,u), V) < \varepsilon.$$

Almost (A,B)-invariant subspaces are of interest to study subspaces invariant under high gain state feedback. Thus, almost (A,B)-invariant subspaces cannot be made invariant under state feedback, so there is no friend, but they can be made almost-invariant in the sense that for every $x \in V$ and any $\varepsilon > 0$ there exists a feedback F such that for all $t \geq 0$

$$\mathrm{dist}(e^{A_F t}x, V) < \varepsilon.$$

Here, for any norm on $\mathbb{R}^n$, the distance of $x \in \mathbb{R}^n$ to a nonvoid subset $A \subset \mathbb{R}^n$ is denoted as

$$\mathrm{dist}(x, A) := \inf_{a \in A} \|x - a\|.$$

In order to derive explicit estimates for the entropy of controlled invariant subspaces it is useful to have explicit parametrizations of the class of all controlled invariant subspaces. This is a difficult task and we refer to e.g. [7],[12] for further information. Special types of subspaces are of special interest here. A controlled invariant subspace V is called coasting, if $V \cap \mathrm{Im} B = \{0\}$. Equivalently, $\{0\}$ is the largest controllability subspace contained in V. Any controlled invariant subspace of a controllable single-input system is coasting. Given an (A, B)–invariant subspace V and a friend $F \in \mathbb{R}^{m \times n}$, then V is A_F-invariant. The restriction $(\bar{A}, \bar{B})$ and co-restriction $(\tilde{A}, \tilde{B})$, respectively, then are defined as

$$\begin{aligned}
(\bar{A}, \bar{B}) &= (A_F|V, B|B^{-1}V), \\
(\tilde{A}, \tilde{B}) &= (A_F : \mathbb{R}^n/V \to \mathbb{R}^n/V, \pi \circ B : \mathbb{R}^m/B^{-1}V \to \mathbb{R}^n/V).
\end{aligned}$$

Note, that the co-restriction $(\tilde{A}, \tilde{B})$ is controllable, whenever (A, B) is controllable, while the restriction is controllable only for a controllability subspace. Note also, that $\bar{B}$ and $\tilde{B}$ are both full column rank if B has full column rank. Of course, the co-restriction may well depend upon the choice of a friend F, so there are in fact many possible co-restrictions and not just one. However, the controllability indices of the co-restrictions are all the same. It is thus a remarkable but simple fact, that for any coasting subspace V, the co-restriction is uniquely defined and independent of F.

If V is an (A, B)-invariant subspace that is coasting, then there is some a-priori information about the dimensions of the bounding subspaces $\langle A\,|V\,\rangle$ and $\ker(A; V)$ that determine the entropy bound (4.1). Here, $\ker(A; V)$ is defined as the largest invariant subspace that is contained in V, while $\ker(A; V)$ denotes the smallest invariant subspace containing V. Generically, one expects $\langle A\,|V\,\rangle = \mathbb{R}^n$ and $\ker(A; V) = \{0\}$, but one can be more specific. For simplicity, we focus on the single input case, i.e. $m = 1$.

LEMMA 2.1. *Let (A, b) be controllable and let V be any (A, b)-invariant subspace that is not A-invariant. Then every A-invariant subspace $W \supset V$ satisfies $W = \mathbb{R}^n$. Any generic (A, b)-invariant subspace satisfies* $\ker(A; V) = \{0\}$ *and* $\langle A\,|V\,\rangle = \mathbb{R}^n$.

Proof. Note, that in the single input case only, every (A, b)-invariant subspace is automatically coasting. By duality, it suffices to show for single-output systems (c, A) that $\langle A\,|V\,\rangle = \mathbb{R}^n$, for any tight (c, A)-invariant subspace V that is not A-invariant. In order to show this we apply the theory of polynomial models; see Fuhrmann [5], Fuhrmann and Willems [6]. Let $q(z) = \det(zI - A)$ denote the characteristic polynomial and X_q denote the associated polynomial model. Thus, X_q denotes the set of polynomials of degree $< n$ with the module structure given by multiplication modulo q. The (tight) conditioned invariant subspaces of codimension d then uniquely correspond to the intersection

$$V = X_q \cap t(z)\mathbb{R}[z], \tag{2.2}$$

via a unique monic polynomial t of degree d; see Fuhrmann and Helmke [7]. Let V_* (V^*) denote the largest (smallest) shift invariant subspace of X_d contained in V (containing V). In this polynomial framework, the A invariant subspaces are of the form $q_1X_{q_2} \subset X_q$, for any factorization $q = q_1q_2$. Thus $q_1X_{q_2} \subset X_q \cap t(z)\mathbb{R}[z]$ if and only if t divides q_1. In particular, t must then divide q, i.e. V must be an invariant subspace. Now assume that V is not an invariant subspace (for the shift), i.e. t does not divide q. Then V does not contain any nontrivial invariant subspace. Applying duality, this implies the lemma. But more can be said. Factor $t = q_1a$ with q_1 a polynomial of degree r dividing q, $q = q_1q_2$, and a a polynomial that is coprime to q. Then $q_1X_{q_2}$ is the smallest invariant subspace containing V. The codimension of this subspace is thus deg gcd $\{q, t\}$. In particular, if q and t are coprime, then $V_* = \{0\}$ and $V^* = X_q$. □

3. Entropy for controlled invariant subspaces. In this section we give two different, but closely related definitions for entropy of a linear subspace $V \subset \mathbb{R}^n$. Our first definitions is a suitable adaptation of the well-known topological entropy of linear differential equations [2].

3.1. Subspace entropy of flows. Let V be a linear subspace of $\mathbb{R}^n$. For a linear map $A : \mathbb{R}^n \to \mathbb{R}^n$, let $\Phi(t, x) = e^{At}x,\ t \in \mathbb{R}_0^+,\ x \in \mathbb{R}^n$, be the induced semiflow. For any compact subset $K \subset V$ and for given $T, \varepsilon > 0$ we call $R \subset K$ a $(T, \varepsilon, K, V, \Phi)$-spanning set, if for all $x \in K$ there exists $y \in R$ with

$$\max_{0\le t\le T} \operatorname{dist}(e^{tA}(x - y), V) < \varepsilon. \tag{3.1}$$

Let $r(T, \varepsilon, K, V, \Phi)$ denote the minimal cardinality of a $(T, \varepsilon, K, V, \Phi)$-spanning set. If no finite $(T, \varepsilon, K, V, \Phi)$-spanning set exists, we set $r(T, \varepsilon, K, V, \Phi) = \infty$. Similarly, we call $S \subset K$ a $(T, \varepsilon, K, V, \Phi)$-separated set, if for all $x \neq y$ in S

$$\max_{0\le t\le T} \operatorname{dist}(e^{tA}(x - y), V) \ge \varepsilon.$$

The maximal cardinality of such a set is denoted by $s(T, \varepsilon, K, V, \Phi)$. Note that the points x in R (and in S) will, in general, not lead to solutions $e^{At}x$ which remain in the ε-neighborhood of V.

Definition 3.1. *Let A be a linear map A on $\mathbb{R}^n$ with associated semiflow Φ and consider a subspace V of $\mathbb{R}^d$. For a compact subset $K \subset V$, we set*

$$h_{\text{span}}(\varepsilon, K, V, \Phi) := \limsup_{T\to\infty} \frac{1}{T} \ln r(T, \varepsilon, K, V, \Phi),$$
$$h_{\text{span}}(K, V, \Phi) := \lim_{\varepsilon\searrow 0} h_{\text{span}}(\varepsilon, K, V, \Phi),$$

and define the entropy of V with respect to Φ by

$$h(V, \Phi) := \sup_K h_{\text{span}}(K, V, \Phi),$$

where the supremum is taken over all compact subsets $K \subset V$.

Analogously, an entropy of V can be defined via minimal separated sets.

As usual in the context of topological entropy, one sees that, by monotonicity, the limits for $\varepsilon \searrow 0$ exist. Since all norms on a finite dimensional vector space are equivalent, the entropy does not depend on the norm used in (3.1). Furthermore, the definitions via separated and spanning sets coincide, which easily follows from the next proposition (cf. Robinson [11, Lemma VIII.1.10]).

PROPOSITION 3.2. *Let $K \subset V$ be compact and fix $T, \varepsilon > 0$. Then*

$$s(T, 2\varepsilon, K, V, \Phi) \leq r(T, \varepsilon, K, V, \Phi) \leq s(T, \varepsilon, K, V, \Phi).$$

Proof. Let $S \subset K$ be a maximal $(T, \varepsilon, K, V, \Phi)$-separated set and let $x \in K$. By maximality of S, there is some $y \in S$ such that

$$\max_{0 \leq t \leq T} \operatorname{dist}(e^{tA}(x - y), V) < \varepsilon.$$

Therefore S is $(T, \varepsilon, K, V, \Phi)$-spanning showing the second inequality. For the first one, consider a maximal $(T, 2\varepsilon, K, V, \Phi)$-separated set S and a minimal $(T, \varepsilon, K, V, \Phi)$-spanning set R. We define a map $H : S \to R$ in the following way: For $x \in S$ there is $y := H(x) \in R$ with $\operatorname{dist}(e^{tA}(x - y), V) < \varepsilon$ for all $t \in [0, T]$. If $H(x_1) = H(x_2) = y$, then

$$\begin{aligned}
&\max_{0 \leq t \leq T} \operatorname{dist}(e^{tA}(x_1 - x_2), V) \\
&\leq \max_{0 \leq t \leq T} \operatorname{dist}(e^{tA}(x_1 - y), V) + \max_{0 \leq t \leq T} \operatorname{dist}(e^{tA}(x_2 - y), V) < 2\varepsilon.
\end{aligned}$$

Thus $x_1 = x_2$ follows. This shows that H is injective, and hence $r(T, \varepsilon, K, V, \Phi) \geq s(T, 2\varepsilon, K, V, \Phi)$. □

Although this will not play any role in the sequel, we describe the behavior of this entropy notion under a special semiconjugacy.

PROPOSITION 3.3. *Let W be an A-invariant subspace for a linear map A on $\mathbb{R}^n$. Then, for a subspace V of $\mathbb{R}^n$ the entropies of the induced flows $\Phi(t, x) = e^{At}x$ on $\mathbb{R}^n$ and $\hat{\bar{\Phi}}(t, \bar{x})$ on the quotient space $\mathbb{R}^n/W$, respectively, satisfy*

$$h(V, \Phi) \geq h(V/W, \bar{\Phi}).$$

Proof. Let $K \subset V$ be compact and for $T, \varepsilon > 0$ consider a $(T, \varepsilon, K, V, \Phi)$-spanning set $R \subset K$. Denote the projection of $\mathbb{R}^n$ to $\mathbb{R}^n/W$ by π, hence $\pi V = V/W$. Then the set πR is a $(T, \varepsilon, \pi K, \pi V, \bar{\Phi})$-spanning set. In fact, let $R = \{x_1, ..., x_\ell\}$ and consider $\pi x \in \pi K$ for some element $x \in K$. Then there exists $x_j \in R$ with

$$\max_{0 \leq t \leq T} \operatorname{dist}(e^{tA}(x - x_j), V) < \varepsilon.$$

Denoting the map induced by A on $\mathbb{R}^n/W$ by $\bar{A}$ one finds for all $t \in [0, T]$

$$\begin{aligned}
\operatorname{dist}(e^{t\bar{A}}(\pi x - \pi x_j), \pi V) &= \inf_{z \in V} \left\| e^{t\bar{A}}(\pi x - \pi x_j) - \pi z \right\| \\
&= \inf_{z \in V, w \in W} \left\| e^{tA}(x - x_j) - z - w \right\| \\
&\leq \operatorname{dist}(e^{tA}(x - x_j), V) \\
&< \varepsilon.
\end{aligned}$$

It follows that the minimal cardinality of a $(T, \varepsilon, K, V, \Phi)$-spanning set is greater than or equal to the minimal cardinality of a $(T, \varepsilon, \pi K, \pi V, \bar{\Phi})$-spanning set. Taking the limit superior for $T \to \infty$, letting ε tend to 0 and, finally, taking the supremum over all compact $K \subset V$ yields the assertion. □

3.2. Entropy for almost (A, B)-invariant subspaces. We now introduce the invariance entropy for almost (A, B)-invariant subspaces of linear control system (2.1) on $\mathbb{R}^n$. In the following, we consider a fixed almost (A, B)-invariant subspace V of $\mathbb{R}^n$ with $\dim V = d$. Furthermore, we admit arbitrary continuous controls in the space $C(\mathbb{R}, \mathbb{R}^m)$ of continuous functions $u : \mathbb{R} \to \mathbb{R}^m$.

DEFINITION 3.4. *For a compact subset $K \subset V$ and for given $T, \varepsilon > 0$ we call a set $\mathcal{R} \subset C(\mathbb{R}, \mathbb{R}^m)$ of control functions a (T, ε, K, V)-spanning set if for all $x_0 \in K$ there is $u \in \mathcal{R}$ with*

$$\operatorname{dist}(\varphi(t, x_0, u), V) < \varepsilon \text{ for all } t \in [0, T].$$

By $r_{\text{inv}}(T, \varepsilon, K, V)$ we denote the minimal cardinality of a (T, ε, K, V)-spanning set. If no finite (T, ε, K, V)-spanning set exists, we set $r_{\text{inv}}(T, \varepsilon, K, V) = \infty$.

In other words: we require for a (T, ε, K, V)-spanning set $\mathcal{R}$ that for every initial value in K, there is a control in $\mathcal{R}$ such that up to time T the trajectory remains in the ε-neighborhood of V.

We note that the definition above differs from earlier ones used for invariance entropy (cf. [4, 3]) by the fact, that the set V whose invariance is studied here, is not compact.

REMARK 3.5. *Let $\varepsilon, T > 0$. By almost (A, B)-invariance of V there exists for every $x \in K$ a control function u with $\operatorname{dist}(\varphi(t, x, u), V) < \varepsilon$ for all $t \geq 0$. Hence, using continuous dependence on initial values and compactness of K, there exist finitely many controls $u_1, ..., u_r$ such that for every $x \in K$ there is u_j with $\varphi(t, x, u_j) \in N_\varepsilon(V)$ for all $t \in [0, T]$. Hence $r_{\text{inv}}(T, \varepsilon, K, V) < \infty$. It seems, that the class of almost (A, B)-invariant subspaces is the largest class of subspaces for which this is true.*

Now we consider the exponential growth rate of $r_{\text{inv}}(T, \varepsilon, K, V)$ for $T \to \infty$ and let $\varepsilon \to 0$.

DEFINITION 3.6. *Let V be an almost (A, B)-invariant subspace. Then, for a compact subset $K \subset V$, the invariance entropy $h_{\text{inv}}(K, V)$ is defined by*

$$h_{\text{inv}}(\varepsilon, K, V) := \limsup_{T \to \infty} \frac{1}{T} \ln r_{\text{inv}}(T, \varepsilon, K, V), \;\; h_{\text{inv}}(K, V) := \lim_{\varepsilon \searrow 0} h_{\text{inv}}(\varepsilon, K, V).$$

Finally, the invariance entropy of V is defined by

$$h_{\text{inv}}(V; A, B) := \sup\nolimits_K h_{\text{inv}}(K, V),$$

where the supremum is taken over all compact subsets $K \subset V$.

In the sequel, we will always use for a given underlying system (A, B) the shorthand notation $h_{\text{inv}}(V)$ for $h_{\text{inv}}(V; A, B)$. Note that $h_{\text{inv}}(\varepsilon_1, K, V) \leq h_{\text{inv}}(\varepsilon_2, K, V)$ for $\varepsilon_2 \leq \varepsilon_1$. Hence the limit for $\varepsilon \to 0$ exists (it might be infinite.) Since all norms on finite dimensional vector spaces are equivalent, the invariance entropy of V is independent of the chosen norm. We will show later that every almost (A, B)-invariant subspace has finite invariance entropy. It is clear by inspection, that both the invariance entropy $h_{\text{inv}}(V)$ and the subspace entropy $h(V, \Phi)$ are invariant under state space similarity; i.e. $h_{\text{inv}}(SV; SAS^{-1}, SB) = h_{\text{inv}}(V; A, B)$.

The following theorem shows that the entropy of an almost invariant (A, B)-subspace V can be characterized by the entropy of V for the corresponding uncontrolled system $\dot{x} = Ax$. This result will be useful in order to compute entropy bounds.

THEOREM 3.7. *Let V be an almost (A,B)-invariant subspace for system (2.1) and consider the entropies $h_{\mathrm{inv}}(V)$ and $h(V,\Phi)$ of V with respect to control system (2.1) and to the linear semiflow $\Phi(t,x)=e^{At}x$, respectively. Then*

$$h_{\mathrm{inv}}(V)=h(V,\Phi).$$

Proof. (i) Let $K\subset V$ be compact, and fix $T,\varepsilon>0$. Consider a (T,ε,K,V)-spanning set $\mathcal{R}=\{u_1,...,u_\ell\}$ of controls with minimal cardinality $r_{\mathrm{inv}}(T,\varepsilon,K,V)$. This means that for every $x\in K$ there is u_j with $\operatorname{dist}(\varphi(t,x,u_j),V)<\varepsilon$ for all $t\in[0,T]$. By minimality, we can, for every u_j, pick $x_j\in K$ with $\operatorname{dist}(\varphi(t,x_j,u_j),V)<\varepsilon$ for all $t\in[0,T]$. Then, using linearity, one finds for all $x\in K$ a control u_j and a point $x_j\in K$ such that for all $t\in[0,T]$

$$\operatorname{dist}(e^{At}x-e^{At}x_j,V)=\operatorname{dist}(\varphi(t,x,u_j)-\varphi(t,x_j,u_j),V)<2\varepsilon.$$

This shows that the points x_j form a $(T,2\varepsilon,K,V,\Phi)$-spanning set, and hence

$$r_{\mathrm{inv}}(T,\varepsilon,K,V)\geq r(T,2\varepsilon,K,V,\Phi).$$

Letting T tend to infinity, then $\varepsilon\to 0$ and, finally, taking the supremum over all compact subsets $K\subset V$, one obtains $h_{\mathrm{inv}}(V)\geq h(V,\Phi)$.

(ii) For the converse inequality, let K be a compact subset of V and $T,\varepsilon>0$. Let $E\subset K$ a maximal (T,ε,K,V)-separated set with respect to the semiflow Φ, say $E=\{y_1,\dots,y_s\}$ with $s=s(T,\varepsilon,K,\ \Phi)$. Then E is also (T,ε,K)-spanning which means that for all $x\in K$ there is $j\in\{1,\dots,s\}$ with

$$\max_{t\in[0,T]}\operatorname{dist}(e^{At}x-e^{At}y_j,V)=\max_{t\in[0,T]}\inf_{z\in V}\left\|e^{At}x-e^{At}y_j-z\right\|<\varepsilon.$$

Since V is almost (A,B)-invariant, we can assign to each $y_j, j\in\{1,\dots,s\}$, a control function $u_j\in C(\mathbb{R},\mathbb{R}^m)$ such that $\varphi(\mathbb{R}_0^+,y_j,u_j)\subset V$. Let $\mathcal{R}:=\{u_1,\dots,u_s\}\subset C(\mathbb{R},\mathbb{R}^m)$. By linearity one has $\varphi(t,x,u)-\varphi(t,y,u)=e^{At}x-e^{At}y$ for all $t\geq 0$, $x,y\in\mathbb{R}^n$ and $u\in C(\mathbb{R},\mathbb{R}^m)$. We obtain that for every $x\in K$ there is j such that

$$\begin{aligned}&\max_{t\in[0,T]}\operatorname{dist}(\varphi(t,x,u_j)-\varphi(t,y_j,u_j),V)\\&=\max_{t\in[0,T]}\operatorname{dist}(e^{At}x-e^{At}y_j,V)\\&<\varepsilon.\end{aligned}$$

Since $\operatorname{dist}(\varphi(t,y_j,u_j),V)<\varepsilon$ for $t\in[0,T]$, there is $z_1\in V$ with $\|\varphi(t,y_j,u_j)-z_1\|<\varepsilon$ and hence

$$\begin{aligned}\operatorname{dist}(\varphi(t,x,u_j),V)&=\inf_{z\in V}\|\varphi(t,x,u_j)-z\|\\&\leq\inf_{z\in V}\|\varphi(t,x,u_j)-\varphi(t,y_j,u_j)+z_1-z\|+\|\varphi(t,y_j,u_j)-z_1\|\\&<\inf_{z\in V}\|\varphi(t,x,u_j)-\varphi(t,y_j,u_j)-z\|+\varepsilon\\&<2\varepsilon.\end{aligned}$$

This implies that for all $x\in K$ there is $u_j\in\mathcal{R}$ such that

$$\max_{t\in[0,T]}\operatorname{dist}(\varphi(t,x,u_j),V)<2\varepsilon.$$

Hence $\mathcal{R}$ is $(T,2\varepsilon,K,V)$-spanning and it follows that

$$r_{\mathrm{inv}}(T,2\varepsilon,K,V)\leq s(T,\varepsilon,K,\Phi)\text{ for all }T,\varepsilon>0,$$

and consequently $h_{\mathrm{inv}}(K,V)\leq h_{\mathrm{sep}}(K,V,\Phi)\leq h(V,\Phi)$. $\square$

4. Entropy Bounds. A general upper bound for entropy of almost (A,B)-invariant subspaces of control system (2.1) is provided. Using the characterization via the subspace entropy of the associated autonomous flow, sharper upper bounds are obtained in special cases.

THEOREM 4.1.

(i) Let $V \subset \mathbb{R}^n$ be any subspace and consider the smallest A-invariant subspace $\langle A\,|V\,\rangle$ containing V and the largest A-invariant subspace $\ker(A;V)$ contained in V. Then the subspace entropy of the flow $\Phi(t,x) = e^{tA}x$ is finite and satisfies

$$h(V,\Phi) \le \max\left(0, \sum \operatorname{Re}\lambda_i\right), \tag{4.1}$$

where summation is over the eigenvalues $\lambda_1,...,\lambda_n$ with $\operatorname{Re}\lambda_i > 0$ of the map induced by A on the quotient space $\langle A\,|V\,\rangle \,/\, \ker(A;V)$.

(ii) The invariance entropy of any almost (A,B)-invariant subspace V for system (2.1) is finite and satisfies the inequality (4.1)

Proof. Note first, that by A-invariance the restriction $A|_{\langle A|V\,\rangle}$ as well as the induced map on $\langle A\,|V\,\rangle \,/\, \ker(A;V)$ are well-defined. Moreover, for any subspace V of the state space there exists a matrix B such that V is an (A,B)-invariant subspace. Thus it suffices to show inequality (4.1) for any almost (A,B)-invariant subspace. It suffices to show that (4.1) is satisfied for $h_{\text{inv}}(K,V)$, where K is an arbitrary compact subset of V. Our proof then depends on a topological entropy inequality for linear maps due to R. Bowen, see the appendix.

Consider the linear semiflow $\Phi(t,x) = e^{At}x$, $\Phi : \mathbb{R}_0^+ \times \mathbb{R}^n \to \mathbb{R}^n$ with time-one map $\Phi_1 = \Phi(1,\cdot)$. By Appendix, Corollary 5.3

$$h_{\text{top}}(\Phi) = h_{\text{top}}(\Phi_1) = \max\left(0, \sum_{i:\ \operatorname{Re}\lambda_i > 0} \operatorname{Re}\lambda_i\right).$$

Analogously, the restriction of Φ_1 to $\langle A\,|V\,\rangle$ has topological entropy given by the sum of the positive real parts of eigenvalues of $\Phi|_{\langle A|V\,\rangle}$. Therefore, the time-one map of the induced flow $\hat{\Phi}$ on $\langle A\,|V\,\rangle \,/\, \ker(A;V)$ has topological entropy given by the right hand side of (4.1); this also coincides with the topological entropy of $\hat{\Phi}$.

Let K be a compact subset of V and π denote the projection of $\langle A\,|V\,\rangle$ to the quotient space $\langle A\,|V\,\rangle \,/\, \ker(A;V)$. Thus, the set πK is compact. Let $T, \varepsilon > 0$ be given and denote by $E \subset \pi K$ a maximal $(T,\varepsilon,\pi K)$-separated set with respect to the semiflow $\hat{\Phi}$ on $\langle A\,|V\,\rangle \,/\, \ker(A;V)$, say $E = \{\pi y_1, \ldots, \pi y_\ell\}$ with $y_j \in K$ and $\ell = s(T,\varepsilon,\pi K, \hat{\Phi})$. Then E is also $(T,\varepsilon,\pi K)$-spanning which means that for all $x \in K$ there is $j \in \{1,\ldots,\ell\}$ with

$$\max_{t\in[0,T]} \operatorname{dist}(e^{At}x - e^{At}y_j, \ker(A;V)) = \max_{t\in[0,T]} \inf_{z\in\ker(A;V)} \left\| e^{At}x - e^{At}y_j - z \right\| < \varepsilon.$$

Since V is almost (A,B)-invariant, we can assign to each y_j, $j \in \{1,\ldots,\ell\}$, a control function $u_j \in C(\mathbb{R},\mathbb{R}^m)$ such that $\operatorname{dist}(\varphi(t,y_j,u_j),V) < \varepsilon$ for all $t \ge 0$. Let $\mathcal{R} := \{u_1,\ldots,u_\ell\} \subset C(\mathbb{R},\mathbb{R}^m)$. By linearity one has $\varphi(t,x,u) - \varphi(t,y,u) = e^{At}x - e^{At}y$ for all $t \ge 0$, $x,y \in \mathbb{R}^n$ and $u \in C(\mathbb{R},\mathbb{R}^m)$. We obtain that for every $x \in K$ there is j

such that

$$
\begin{aligned}
&\max_{t\in[0,T]} \operatorname{dist}(\varphi(t,x,u_j)-\varphi(t,y_j,u_j),V)\\
&\le \max_{t\in[0,T]} \operatorname{dist}(\varphi(t,x,u_j)-\varphi(t,y_j,u_j),\ker(A;V))\\
&= \max_{t\in[0,T]} \operatorname{dist}(e^{At}x-e^{At}y_j,\ker(A;V))\\
&<\varepsilon.
\end{aligned}
$$

Since $\operatorname{dist}(\varphi(t,y_j,u_j),V)<\varepsilon$ for $t\in[0,T]$, there is $z_1\in V$ with

$$\|\varphi(t,y_j,u_j)-z_1\|<\varepsilon$$

and hence, using that V is a linear subspace, one finds

$$
\begin{aligned}
\operatorname{dist}(\varphi(t,x,u_j),V) &= \inf_{z\in V}\|\varphi(t,x,u_j)-z\|\\
&\le \inf_{z\in V}\|\varphi(t,x,u_j)-\varphi(t,y_j,u_j)+z_1-z\|+\|\varphi(t,y_j,u_j)-z_1\|\\
&< \inf_{z\in V}\|\varphi(t,x,u_j)-\varphi(t,y_j,u_j)-z\|+\varepsilon\\
&<2\varepsilon.
\end{aligned}
$$

This implies that for all $x\in K$ there is $u_j\in\mathcal{R}$ such that $\operatorname{dist}(\varphi(t,x,u_j),V)<2\varepsilon$ for all $t\in[0,T]$. Hence $\mathcal{R}$ is $(T,2\varepsilon,K,V)$-spanning and it follows that

$$r_{\text{inv}}(T,2\varepsilon,K,V)\le s(T,\varepsilon,\pi K,\hat{\Phi}) \text{ for all } T,\varepsilon>0,$$

and consequently $h_{\text{inv}}(K,V)\le h_{\text{sep}}(\pi K,\hat{\Phi})=h_{\text{top}}(\pi K,\hat{\Phi})\le h_{\text{top}}(\hat{\Phi})$.□

The above bound is rather conservative and can be improved in several cases. We therefore turn to the computation of sharper bounds of $h(V,\Phi)$ under suitable genericity conditions on the almost controlled invariant subspace V. Here one will expect that starting in a neighborhood of the origin in V, the maximal real parts of eigenvalues determine the behavior. This will be made precise below.

We begin with a few lemmas. In the sequel, $e_1,\cdots,e_n$ denotes the standard basis vectors of $\mathbb{R}^n$. For a real diagonalizable matrix A, order the eigenvalues of A such that

$$\lambda_1\ge\lambda_2\ge\ldots\ge\lambda_n.$$

Lemma 4.2. *Let $A\in\mathbb{R}^{n\times n}$ be diagonalizable and consider a d-dimensional subspace $V\subset\mathbb{R}^n$, such that the transversality condition $V\cap W=\{0\}$ holds for any $(n-d)$-dimensional A-invariant subspace $W\subset\mathbb{R}^n$. We can write the Jordan representation J of A as*

$$J=\operatorname{diag}(\lambda_1,\ldots,\lambda_n),$$

and we abbreviate

$$\Lambda_1:=\operatorname{diag}(\lambda_1,\ldots,\lambda_d),\ \text{and}\ \Lambda_2:=\operatorname{diag}(\lambda_{d+1},\ldots,\lambda_n),. \tag{4.2}$$

Then there exist $S\in GL_n(\mathbb{R})$ and $G\in\mathbb{R}^{(n-d)\times d}$ with $V=\{Se_1,\cdots,Se_d\}$, and

$$S^{-1}AS=\begin{bmatrix}\Lambda_1 & 0\\ G & \Lambda_2\end{bmatrix}. \tag{4.3}$$

Proof. Let $w_1, ..., w_n$ be a corresponding basis of eigenvectors and denote

$$V_1 := \langle w_1, ..., w_d \rangle \text{ and } W = \langle w_{d+1}, ..., w_n \rangle .$$

By assumption on V, we have $V \cap W = \{0\}$ and therefore the canonical projection map$\pi : \mathbb{R}^n \to V_1$ along W maps V isomorphically onto V_1. Choose any basis $v_1, \cdots, v_d$ of V and extend it to a basis $S_1 = (v_1, \cdots, v_d, w_{d+1}, \cdots, w_n)$. Then

$$S_1^{-1} A S_1 = \begin{bmatrix} \Gamma_1 & 0 \\ A_2 & \Lambda_2 \end{bmatrix},$$

and therefore Γ_1 has the same eigenvalues as Λ_1. Finally, we can transform Γ_1 to Jordan normal form by a matrix S_2. Then conjugation with the matrix

$$S_1 \cdot \begin{bmatrix} S_2 & 0 \\ 0 & I \end{bmatrix}$$

leads to (4.3).□

In the situation as above, by invariance of the problem under similarity, we can assume without loss of generality, that

$$A = \begin{bmatrix} \Lambda_1 & 0 \\ G & \Lambda_2 \end{bmatrix}, \; V = \mathbb{R}^d \times \{0\}. \tag{4.4}$$

Note that for any vector $z = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^n = \mathbb{R}^d \times \mathbb{R}^{n-d}$ we have $\operatorname{dist}(z, V) = \|y\|$. For $x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} \in V$ we compute

$$e^{tA}(x-y) = \begin{bmatrix} e^{t\Gamma_1} & 0 \\ M(t) & e^{t\Lambda_2} \end{bmatrix} \begin{pmatrix} x_1 - y_1 \\ 0 \end{pmatrix} = \begin{bmatrix} e^{t\Gamma_1}(x_1 - y_1) \\ M(t)(x_1 - y_1) \end{bmatrix}$$

and

$$\operatorname{dist}(e^{tA}(x-y), V) = \|M(t)(x_1 - y_1)\| .$$

The function $M(t)$ is the unique solution to the linear differential equation $\dot{M} = \Lambda_2 M + G e^{t\Lambda_1}$ with initial condition $M(0) = 0$ and therefore

$$M(t) = e^{t\Lambda_2} \int_0^t e^{-s\Lambda_2} G e^{s\Lambda_1} ds.$$

The formula for $M(t)$ shows that for diagonal Λ_2 and Γ_1 one finds, with $e_j = j$th standard basis vector and $g_j = Ge_j \in \mathbb{R}^{n-d}$, $j = 1, ..., d$, for the jth column of $M(t)$

$$\begin{aligned} M(t)e_j &= e^{t\Lambda_2} \int_0^t e^{-s\Lambda_2} G e^{s\Lambda_1} ds \; e_j = e^{t\Lambda_2} \int_0^t e^{-s\Lambda_2} G e^{s\lambda_j} e_j ds \\ &= e^{t\Lambda_2} \int_0^t e^{s(\lambda_j I_{n-d} - \Lambda_2)} ds \; g_j. \end{aligned}$$

Let for $k = 1, ..., n-d, \; j = 1, ..., d$

$$\alpha_{kj}(t) := \begin{cases} \frac{1}{\lambda_j - \lambda_{d+k}} \left[1 - e^{t(\lambda_{d+k} - \lambda_j)}\right] & \text{for} \quad \lambda_j > \lambda_{d+k} \\ t & \text{for} \quad \lambda_j = \lambda_{d+k} \end{cases} .$$

They satisfy for $t > 0$ the inequalities

$$0 < \alpha_{kj}(t) \leq \begin{cases} (\lambda_j - \lambda_{d+k})^{-1} & \text{for} \quad \lambda_j > \lambda_{d+k} \\ t & \text{for} \quad \lambda_j = \lambda_{d+k} \end{cases}.$$

Furthermore, for $t \geq 1$ one has

$$\alpha_{kj}(t) \geq \begin{cases} \frac{1}{\lambda_j - \lambda_{d+k}} [1 - e^{\lambda_{d+k} - \lambda_j}] & \text{for} \quad \lambda_j > \lambda_{d+k} \\ 1 & \text{for} \quad \lambda_j = \lambda_{d+k} \end{cases}$$

(recall that $\lambda_j \geq \lambda_{d+k}$ for all $j = 1, ..., d$ and all $k = 1, ..., n-d$.)

One computes that

$$M(t) e_j = e^{t\lambda_j} \operatorname{diag}[\alpha_{1j}(t), ..., \alpha_{n-d,j}(t)] \ g_j, \ \ j = 1, ..., d. \tag{4.5}$$

Proposition 4.3. *Let $A \in \mathbb{R}^{n \times n}$ be diagonalizable and consider a d-dimensional subspace $V \subset \mathbb{R}^n$ which satisfies $V \cap W = \{0\}$ for any $(n-d)$-dimensional A-invariant subspace $W \subset \mathbb{R}^n$. Assume that the eigenvalues of A satisfy*

$$\lambda_1 \geq ... \geq \lambda_d > \lambda_{d+1} \geq ... \geq \lambda_n.$$

Then the entropy $h(V, \Phi)$ of V with respect to the linear semiflow $\Phi(t, x) = e^{At}x$ is bounded above by the topological entropy of the semiflow $\Phi_1(t, x) = e^{\Lambda_1 t}x$ where $\Lambda_1 = \operatorname{diag}[\lambda_1, ..., \lambda_d]$. Thus it satisfies the upper bound

$$h(V, \Phi) \leq \max\left(0, \sum \lambda_i\right), \tag{4.6}$$

where summation is over the positive eigenvalues $\lambda_1, ..., \lambda_d$.

Proof. Let $K \subset V$ be compact. We show that for $T, \varepsilon > 0$ any $(T, \varepsilon, K, \Phi_1)$-spanning set R for the topological entropy of the semiflow Φ_1 is $(T, c\varepsilon, K, V, \Phi)$-spanning for Φ, with

$$c := (\lambda_d - \lambda_{d+1})^{-1} \max_{j=1,\dots,d} \|g_j\|.$$

For every $x \in K$ there is $y \in R$ such that, by formula (4.5),

$$\begin{aligned} \|M(t)(x-y)\| &= \left\| \sum_{j=1}^{d} M(t) e_j (x_j - y_j) \right\| \\ &= \left\| \sum_{j=1}^{d} e^{t\lambda_j} \operatorname{diag}[\alpha_{1j}(t), ..., \alpha_{n-d,j}(t)] \ g_j (x_j - y_j) \right\| \\ &\leq \left| \sum_{j=1}^{d} e^{t\lambda_j} (x_j - y_j) \right| (\lambda_j - \lambda_{d+k})^{-1} \|g_j\| \\ &\leq \left\| e^{t\Lambda_1}(x-y) \right\| (\lambda_d - \lambda_{d+1})^{-1} \max_{j=1,\dots,d} \|g_j\| \\ &\leq c \max_{t \in [0,T]} \left\| e^{t\Lambda_1}(x-y) \right\| \\ &< c\, \varepsilon. \end{aligned}$$

This implies inequality (4.6). $\square$

Proposition 4.3 yields the following estimate for the invariance entropy of almost (A,B)-invariant subspaces. This estimate is sharper than the one provided in Theorem 4.1 (in particular, for low dimensional spaces V.)

THEOREM 4.4. *Consider an almost (A,B)-invariant subspace $V \subset \mathbb{R}^n$ with dimension d and denote the largest A-invariant subspace contained in V by $\ker(A;V)$ and its dimension by ℓ. Suppose that the map $\overline{A}$ induced by A on the quotient space $\mathbb{R}^n/\ker(A;V)$ is diagonalizable and that the eigenvalues satisfy*

$$\lambda_1 \geq \ldots \geq \lambda_{d-\ell} > \lambda_{d+1-\ell} \geq \ldots \geq \lambda_{n-\ell}.$$

Assume further that $V/\ker(A;V)$ intersects trivially any $\overline{A}$-invariant subspace $W \subset \mathbb{R}^n/\ker(A;V)$ of codimension $d-\ell$. Then the invariance entropy of V satisfies the inequality

$$h_{\mathrm{inv}}(V) \leq \max\left(0, \sum \lambda_i\right), \tag{4.7}$$

where summation is over the positive eigenvalues $\lambda_i,\ i \in \{1,\ldots,d-\ell\}$.

Proof. We argue similarly as in the proof of Theorem 4.1, using now the subspace entropy with respect to V instead of the topological entropy.

Let K be a compact subset of V. Then, for the projection π of $\mathbb{R}^n$ to the quotient space $\mathbb{R}^n/\ker(A;V)$, the set πK is compact. Let $T,\varepsilon > 0$ be given and denote by $E \subset \pi K$ a maximal $(T,\varepsilon,\pi K,\pi V,\hat{\Phi})$-separated set with respect to the semiflow $\hat{\Phi}$ on $\mathbb{R}^n/\ker(A;V)$, say $E = \{\pi y_1,\ldots,\pi y_s\}$ with $y_j \in K$ and $s = s(T,\varepsilon,\pi K,\pi V,\hat{\Phi})$. Then E is also $(T,\varepsilon,\pi K,\pi V,\hat{\Phi})$-spanning which means that for all $x \in K$ there is $j \in \{1,\ldots,s\}$ with

$$\max_{t\in[0,T]} \mathrm{dist}(e^{At}x - e^{At}y_j, V + \ker(A;V)) = \max_{t\in[0,T]} \inf_{z\in V} \left\| e^{At}x - e^{At}y_j - z \right\| < \varepsilon.$$

Since V is almost (A,B)-invariant, we can assign to each $y_j, j \in \{1,\ldots,s\}$, a control function $u_j \in C(\mathbb{R},\mathbb{R}^m)$ such that $\mathrm{dist}(\varphi(t,y_j,u_j),V) < \varepsilon$ for all $t \geq 0$. Let $\mathcal{R} := \{u_1,\ldots,u_s\} \subset C(\mathbb{R},\mathbb{R}^m)$. By linearity one has $\varphi(t,x,u) - \varphi(t,y,u) = e^{At}x - e^{At}y$ for all $t \geq 0,\ x,y \in \mathbb{R}^n$ and $u \in C(\mathbb{R},\mathbb{R}^m)$. We obtain that for every $x \in K$ there is j such that

$$\max_{t\in[0,T]} \mathrm{dist}(\varphi(t,x,u_j) - \varphi(t,y_j,u_j),V) = \max_{t\in[0,T]} \mathrm{dist}(e^{At}x - e^{At}y_j,V) < \varepsilon.$$

Since $\mathrm{dist}(\varphi(t,y_j,u_j),V) < \varepsilon$ for $t \in [0,T]$, there is $z_1 \in V$ with

$$\|\varphi(t,y_j,u_j) - z_1\| < \varepsilon$$

and hence, using that V is a linear subspace, one finds

$$\begin{aligned}
\mathrm{dist}(\varphi(t,x,u_j),V) &= \inf_{z\in V} \|\varphi(t,x,u_j) - z\| \\
&\leq \inf_{z\in V} \|\varphi(t,x,u_j) - \varphi(t,y_j,u_j) + z_1 - z\| + \|\varphi(t,y_j,u_j) - z_1\| \\
&< \inf_{z\in V} \|\varphi(t,x,u_j) - \varphi(t,y_j,u_j) - z\| + \varepsilon \\
&< 2\varepsilon.
\end{aligned}$$

This implies that for all $x \in K$ there is $u_j \in \mathcal{R}$ such that

$$\max_{t\in[0,T]} \operatorname{dist}(\varphi(t,x,u_j),V) < 2\varepsilon.$$

Hence $\mathcal{R}$ is $(T,2\varepsilon,K,V)$-spanning and it follows that

$$r_{\text{inv}}(T,2\varepsilon,K,V) \le s(T,\varepsilon,\pi K,\pi V,\hat{\Phi}) \text{ for all } T,\varepsilon > 0,$$

and consequently

$$h_{\text{inv}}(K,V) \le h_{\text{sep}}(\pi K,\pi V,\hat{\Phi}) = h(\pi K,\pi V,\hat{\Phi}) \le h(\pi V,\hat{\Phi}).$$

Since, by assumption, $\pi V = V/\ker(A;V)$ intersects trivially any $\overline{A}$-invariant subspace $W \subset \mathbb{R}^n/\ker(A;V)$ of codimension $d-\ell$. does not contain any nontrivial A-invariant subspace, we can apply Proposition 4.3 in order to prove the assertion. □

We list a few explicit cases in the single-input case in which the hypotheses of Theorem 4.4 are satisfied.

COROLLARY 4.5. *Assume that $(A,b) \in \mathbb{R}^{n\times n} \times \mathbb{R}^n$ is controllable and A is diagonalizable with n distinct real eigenvalues*

$$\lambda_1 > \ldots > \lambda_d > \lambda_{d+1} > \ldots > \lambda_n.$$

Let $\alpha_1,\cdots,\alpha_d$ denote any distinct real numbers that are disjoint from $\lambda_1,\cdots,\lambda_n$. Then

$$V = \operatorname{span}\left((A-\alpha_1 I)^{-1}b,\cdots,(A-\alpha_d I)^{-1}b)\right) \tag{4.8}$$

is an (A,b)-invariant subspace with $\ker(A;V)=\{0\}$ and $< A|V >= \mathbb{R}^n$. The entropy of V satisfies the inequality

$$h_{\text{inv}}(V) \le \max\left(0,\sum_{i=1}^{d}\lambda_i\right), \tag{4.9}$$

Proof. Without loss of generality we can assume that (A,b) is in Jordan canonical form, i.e. $A = \operatorname{diag}(\lambda_1,\cdots,\lambda_n)$ and $b = (1,\cdots,1)^\top$. Thus V coincides with the column span of the $n\times d$ matrix

$$\begin{bmatrix} (\lambda_1-\alpha_1)^{-1} & \cdots & (\lambda_1-\alpha_d)^{-1} \\ \vdots & & \vdots \\ (\lambda_n-\alpha_1)^{-1} & \cdots & (\lambda_n-\alpha_d)^{-1} \end{bmatrix}.$$

For any column v of this matrix, the pair (A,v) is controllable, which implies $< A|V >$ $= \mathbb{R}^n$. Let W denote an arbitrary A-invariant eigenspace of codimension d and assume $W \cap V \neq \{0\}$. Then there exists nonzero real numbers $c_1,\cdots c_d$ such that the rational function

$$\frac{p(\lambda)}{q(\lambda)} := \sum_{i=1}^{d}\frac{c_i}{\lambda-\alpha_i}$$

vanishes at d eigenvalues $\lambda \in \{\lambda_{i_1},\cdots,\lambda_{i_d}\}$. But then $p=0$, as $\deg p < d$ and therefore $c_1 = \cdots = c_d = 0$, which is a contradiction. Hence $W \cap V = \{0\}$ for any A-invariant subspace of codimension d. This implies also $\ker(A; V) = \{0\}$, as any invariant subspace $V_0 \subset V$ can be extended to an $(n-d)$-dimensional invariant subspace W. This shows that V satisfies the assumptions of Theorem 4.4. □

We note that the (A, b)-invariant subspaces constructed in Corollary 4.5 are not of the most general form; however for $d = 1$ they parameterize all one-dimensional controlled invariant subspaces. In the scalar case, it can be shown that the estimate above is sharp.

EXAMPLE 4.6. *Let $d = 1, n = 2$. We can suppose that A has the form (4.4) and we use small letters instead of capital letters. Let $K \subset V = \mathbb{R} \times \{0\}$ be compact. Choose a $(T, \varepsilon, K, V, \Phi)$-spanning set $R \subset K$. Thus for all $x \in K$ there is $y \in R$ such that for all $t \in [0, T]$*

$$\operatorname{dist}\left(e^{tA}\left(\begin{bmatrix} x \\ 0 \end{bmatrix} - \begin{bmatrix} y \\ 0 \end{bmatrix}\right), V\right) = \|m(t)(x-y)\| < \varepsilon.$$

If V is not invariant, one has $g \neq 0$. For $t \geq 1$

$$\begin{aligned} \varepsilon > \|m(t)(x-y)\| &= \left\| e^{t\lambda_2} \int_0^t e^{-s\lambda_2} g e^{s\lambda_1} ds \ (x-y) \right\| \\ &= \begin{cases} \frac{g}{\lambda_1 - \lambda_2} e^{t\lambda_1} [1 - e^{t(\lambda_2 - \lambda_1)}] \|x-y\| & \text{for} \quad \lambda_1 > \lambda_2 \\ e^{t\lambda_1} t\,|g| \ \ \|x-y\| & \text{for} \quad \lambda_1 = \lambda_2 \end{cases} \\ &\geq c e^{t\lambda_1} \|x-y\|, \end{aligned}$$

with a constant $c > 0$ given by

$$c := \begin{cases} \frac{|g|}{\lambda_1 - \lambda_2} [1 - e^{\lambda_2 - \lambda_1}] & \text{for} \quad \lambda_1 > \lambda_2 \\ |g| & \text{for} \quad \lambda_1 = \lambda_2 \end{cases}$$

(recall that $\lambda_1 \geq \lambda_2$.) Hence

$$e^{t\lambda_1} \|x-y\| \leq c^{-1}\varepsilon \text{ for } t \in [1, T].$$

Remark 5.4 shows, that the set R is a spanning set for the topological entropy of the flow $e^{t\lambda_1}, t \geq 0,\ x \in V$. It follows that

$$r(T, c^{-1}\varepsilon, K, V) \geq r_{top}(T, \varepsilon, K, e^{\Lambda_1 \cdot}).$$

Hence $h_{\mathrm{inv}}(V) \geq h_{top}(e^{\Lambda_1 \cdot})$ follows showing that equality holds in (4.9).

EXAMPLE 4.7. *Here we treat the n-dimensional generalization of the above example, i.e. $d = 1$ and $n \geq 2$. Assume further, that $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ and $G = (g_j) \in \mathbb{R}^{(n-1)\times 1}$ is nonzero. Then*

$$\|M(t)(x-y)\| = e^{t\lambda_1} \|x-y\| \|v(t)\|$$

with (we take the 1-norm)

$$\|v(t)\| = \sum_{j=2}^{n} |g_j| \frac{1 - e^{t(\lambda_j - \lambda_1)}}{\lambda_1 - \lambda_j}$$

upper bounded on $[0, \infty)$ by $\bar{c} := \sum_{j=2}^{n} |g_j| \frac{1}{\lambda_1 - \lambda_j}$ and lower bounded on $[1, \infty)$ by $\underline{c} = \sum_{j=2}^{n} |g_j| \frac{1 - e^{\lambda_j - \lambda_1}}{\lambda_1 - \lambda_j}$. Proceeding as in the above example, we conclude that the entropy is given by $\max(0, \lambda_1)$.

5. Appendix: Topological entropy of linear maps. In this section we recall the definition and characterization of topological entropy for linear maps from Bowen [2]). For the reader's convenience, we also provide a proof of the characterization in terms of eigenvalues; this is a special case of the more general result in [2].

We consider a linear map $A : \mathbb{R}^d \to \mathbb{R}^d$. Let $K \subset \mathbb{R}^d$ be a compact set and fix $\varepsilon > 0$ and $n \in \mathbb{N}$. A set $S \subset K$ is called $(n, \varepsilon, K; A)$-separated if for all $x, y \in K$ with $x \neq y$ there is $i \in \{0, 1, ..., n-1\}$ with $\left\|A^i x - A^i y\right\| > \varepsilon$. Denote by $s(n, \varepsilon, K; A)$ the maximal cardinality of an $(n, \varepsilon, K; A)$-separated set. A set $R \subset K$ is called $(n, \varepsilon, K; A)$-spanning if for every $x \in K$ there is $y \in R$ such that for all $i \in \{0, 1, ..., n-1\}$ one has $\left\|A^i x - A^i y\right\| \leq \varepsilon$. Denote by $s(n, \varepsilon, K; A)$ the minimal cardinality of an $(n, \varepsilon, K; A)$-spanning set. Then the topological entropy of A with respect to K is defined as

$$h_{sep}(\varepsilon, K; A) := \limsup_{n \to \infty} \frac{1}{n} \ln s(n, \varepsilon, K; A), \;\; h_{sep}(K; A) := \lim_{\varepsilon \searrow 0} h_{sep}(\varepsilon, K; A),$$

and, finally,

$$h_{top}(A) := \sup\nolimits_K h_{sep}(K; A).$$

It is easily seen that the topological entropy can also be defined via spanning sets, using instead of $h_{sep}(\varepsilon, K; A)$

$$h_{span}(\varepsilon, K; A) := \limsup_{n \to \infty} \frac{1}{n} \ln r(n, \varepsilon, K; A).$$

This follows, since a maximal $(n, \varepsilon, K; A)$-separated set is also $(n, \varepsilon, K; A)$-spanning, which implies $s(n, \varepsilon, K; A) \geq r(n, \varepsilon, K; A)$; furthermore, using the triangle inequality one sees that $s(n, \varepsilon, K; A) \leq r(n, 2\varepsilon, K; A)$. Topological entropy of linear maps can be characterized by the eigenvalues of A.

THEOREM 5.1. *For a linear map $A : \mathbb{R}^d \to \mathbb{R}^d$, the topological entropy is given by*

$$h_{top}(A) = \max\left(0, \sum \log |\lambda_i|\right),$$

where the sum is taken over all eigenvalues $\lambda_i, i = 1, ..., d$, of A with $|\lambda_i| > 1$ (if no eigenvalue λ_i with $|\lambda_i| > 1$ exists, the sum is omitted.)

Proof. Without loss of generality, there is an eigenvalue with absolute value > 1. Decompose $\mathbb{R}^d$ into two subspaces $\mathbb{R}^d = X^+ \oplus X^-$, where X^+ is the sum of all (real) generalized eigenspaces corresponding to eigenvalues with absolute value greater than 1 and X^- is the sum of all eigenspaces corresponding to eigenvalues with absolute value equal to or less than 1. These subspaces are invariant under A, i.e., the restrictions

$$A^+ := A_{|X^+} : X^+ \to X^+ \text{ and } A^- := A_{|X^-} : X^- \to X^-$$

are well defined. Then, using an elementary property of topological entropy (cf. Pollicott and Yuri [10])

$$h_{top}(A) = h_{top}(A^+) + h_{top}(A^-) \geq h_{top}(A^+). \tag{5.1}$$

First we show $h_{top}(A) \geq \sum \log |\lambda_i|$. By (5.1) it suffices to show the estimate for A^+. In other words, we may assume without loss of generality that all eigenvalues of A

have absolute value greater than 1. Let $K \subset \mathbb{R}^d$ be compact with positive (Lebesgue-) measure $\mu(K) > 0$. Then for $n \in \mathbb{N}$

$$\mu(A^n(K)) = |\det A|^n \, \mu(K) = \left(\prod |\lambda|\right)^n \mu(K) = e^{n \sum \log|\lambda|} \mu(K).$$

Let $\varepsilon > 0, n \in \mathbb{N}$, and consider an $(n+1, \varepsilon, K)$-spanning set R of minimal cardinality $r(n+1, \varepsilon, K)$. Then (by the definition of spanning sets) the set $A^n K$ is contained in the union of $r(n+1, \varepsilon, K)$ balls $B(x_j, \varepsilon)$ of radius ε. Each ball has measure $(2\varepsilon)^d$ (take the max-norm). Thus

$$\mu(A^n(K)) \leq r(n+1, \varepsilon, K) \cdot (2\varepsilon)^d.$$

This yields

$$\log r(n+1, \varepsilon, K) \geq \log \mu(A^n(K)) - \log 2\varepsilon^d \geq n \sum \log|\lambda| + \log \mu(K) - \log 2\varepsilon^d.$$

and hence

$$\limsup_{n\to\infty} \frac{1}{n} \log r(n, \varepsilon, K) = \limsup_{n\to\infty} \frac{n}{n+1} \frac{1}{n} \log r(n+1, \varepsilon, K) \geq \sum \log|\lambda|.$$

For the converse inequality, decompose $\mathbb{R}^d$ into linear subspaces which are the sums of the (real) generalized eigenspaces V_i for eigenvalues of equal absolute values $|\lambda_i|$. These subspaces are A-invariant and hence every restriction $A|_{V_i} : V_i \to V_i$ has only eigenvalues of equal absolute value $|\lambda_i|$ and the sum of their algebraic multiplicities equals $d_i := \dim V_i$. Then one finds

$$h_{top}(A) = h_{top}(A|_{V_1}) + \ldots + h_{top}(A|_{V_r}) \leq \sum_i \max(0, d_i \log|\lambda_i|) = \sum \max(0, \log|\lambda_i|),$$

where the inequality follows by Lemma 5.2(ii). □

The following lemma is needed in the proof above.

LEMMA 5.2. *Let A be a linear map A on $\mathbb{R}^d$. Then, for an eigenvalue$\lambda_{\max}$ of A with maximal absolute value, one has*

$$h_{top}(A) \leq \max(0, d \log|\lambda_{\max}|).$$

Proof. We first show

$$h_{top}(A) = \sup\nolimits_K h(A, K) \leq \max(0, d \log\|A\|). \tag{5.2}$$

If $\|A\| \leq 1$, then every $(1, \varepsilon, K)$-spanning set is also (n, ε, K)-spanning for $n \geq 1$. Hence $h_{top}(A) = 0$ and there is nothing to prove. So we may assume $\|A\| > 1$. Let $K \subset \mathbb{R}^d$ be compact. Then there is $N \in \mathbb{N}$ with

$$K \subset [-N, N]^d.$$

For $\delta > 0$ and $M := \lceil \frac{1}{\delta} \rceil \in \mathbb{N}$, every point in $[-N, N]$ has distance less that $\frac{1}{M} \leq \delta$ to one of the $2MN + 1$ points in

$$S := \{x_i = \frac{i}{M}, i = -N, \ldots, N\}.$$

Then, in the max-norm, every point in $K \subset [-N, N]^d$ has distance less than $\frac{1}{M} \leq \delta$ to one of the $(2MN+1)^d$ in the product S^d. Since

$$\left\| A^i x - A^i y \right\| \leq \|A\|^i \|x - y\| ,$$

we see that S^d is an $(n, \delta \|A\|^n)$-spanning set of cardinality $(2MN+1)^d \leq M^d(2N+1)^d = \left\lceil \frac{1}{\delta} \right\rceil^d (2N+1)^d \leq \left(\frac{1}{\delta}+1\right)^d (2N+1)^d$.

Thus for $\varepsilon > 0$ and $\delta := \frac{\varepsilon}{\|A\|^n}$ we find that there is an (n, ε)-spanning set of cardinality equal to or less than

$$\left(\frac{\|A\|^n}{\varepsilon}+1\right)^d (2N+1)^d = \left(\frac{\|A\|^n+\varepsilon}{\varepsilon}\right)^d (2N+1)^d = (\|A\|^n+\varepsilon)^d \varepsilon^{-d} (2N+1)^d.$$

We find that the minimal cardinality $r(n, \varepsilon, K)$ of an (n, ε, K)-spanning set satisfies

$$r(n, \varepsilon, K) \leq (\|A\|^n+\varepsilon)^d \varepsilon^{-d} (2N+1)^d,$$

and hence

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log r(n, \varepsilon, K) &\leq \limsup_{n\to\infty} \frac{1}{n} \left[d \log(\|A\|+\varepsilon) - \log \varepsilon^d + \log(2N+1)^d \right] \\ &= d \log(\|A\|+\varepsilon). \end{aligned}$$

For $\varepsilon \to 0$ assertion (5.2) follows.

By Abramov's Theorem (cf. Pollicott and Yuri [10]) and (5.2), we find for $n \geq 1$

$$\begin{aligned} h_{top}(A) &= \frac{1}{n} h_{top}(A^n) \leq \frac{1}{n} \max(0, d \log \|A^n\|) = \max(0, d \frac{1}{n} \log \|A^n\|) \\ &= \max(0, d \log \|A^n\|^{1/n}). \end{aligned}$$

Finally, by a standard property of matrices, $\|A^n\|^{1/n}$ converges for $n \to \infty$, to the spectral radius of A which equals $|\lambda_{\max}|$. □

For linear flows $\Phi(t, x) = e^{At} x$, $\Phi : \mathbb{R}_0^+ \times \mathbb{R}^n \to \mathbb{R}^n$ the topological entropy $h_{\mathrm{top}}(\Phi)$ is defined analogously as for maps, based on the following notion. For a compact subset $K \subset \mathbb{R}^n$ and $T > 0, \varepsilon > 0$, a set $R \subset K$ is called $(T, \varepsilon, K, \Phi)$-spanning if for every $x \in K$ there is $y \in R$ such that

$$\max_{t\in[0,T]} \|\Phi(t, x) - \Phi(t, y)\| \leq \varepsilon.$$

COROLLARY 5.3. *Consider the linear semiflow $\Phi(t, x) = e^{At} x$, $\Phi : \mathbb{R}_0^+ \times \mathbb{R}^n \to \mathbb{R}^n$. The topological entropy $h_{\mathrm{top}}(\Phi)$ equals the topological entropy of the time-one-map $\Phi_1(x) = e^A x$ and hence*

$$h_{\mathrm{top}}(\Phi) = \max\left(0, \sum \operatorname{Re} \lambda_i\right),$$

where the sum is taken over all eigenvalues λ_i of A with positive real parts $\operatorname{Re} \lambda_i$.

Proof. The topological entropy $h_{\mathrm{top}}(\Phi)$ equals the topological entropy of the time-one-map $\Phi_1(x) = e^A x$ (see e.g. [4, Lemma 2.1]). By Theorem 5.1, the topological entropy of the linear map Φ_1 is given by

$$h_{\mathrm{top}}(\Phi_1) = \max\left(0, \sum \ln |\nu_i|\right),$$

where the sum is taken over all eigenvalues ν_i of e^A with $|\nu_i| > 1$. Since $|\nu_i| = \left|e^{\lambda_i}\right| = e^{\operatorname{Re}\lambda_i}$ for the eigenvalues λ_i of A, we obtain

$$h_{\text{top}}(\Phi) = h_{\text{top}}(\Phi_1) = \max\left(0, \sum \operatorname{Re}\lambda_i\right),$$

where the sum is taken over all eigenvalues λ_i of A with $\operatorname{Re}\lambda_i > 0$. □

REMARK 5.4. *An easy consequence of this result is, that one also may define (T, ε, K)-spanning sets R by requiring for $T > 1$, $\varepsilon > 0$ that for every $x \in K$ there is $y \in R$ with $\|\Phi(t,x) - \Phi(t,y)\| \leq \varepsilon$ for all $t \in [1,T]$. This leads to the same topological entropy. The same is true for separated sets.*

REFERENCES

[1] G. BASILE AND G. MORRO, *Controlled and conditioned invariant subspaces in linear system theory*, J. Optim. Theory Appl., 3 (1969), pp. 306–315.

[2] R. BOWEN, *Entropy for group endomorphisms and homogeneous spaces*, Trans. Amer. Math. Soc., 153 (1971).

[3] F. COLONIUS AND C. KAWAN, *Asymptotic invariance entropy and outputs*. submitted, 2009.

[4] ———, *Invariance entropy for control systems*, SIAM J. Control Optim., 48 (2009), pp. 1701–1721.

[5] P.A. FUHRMANN, *Linear Operators and Systems in Hilbert Space*, McGraw-Hill, New York, 1981.

[6] P.A. FUHRMANN AND J.C. WILLEMS *A study of (A, B)-invariant subspaces via polynomial models*, Int. J. Control, 31 (1980), pp. 467–494.

[7] P.A. FUHRMANN AND U. HELMKE *On the parametrization of conditioned invariant subspaces and observer theory*, Linear Algebra and its Appl., 332–334 (2001), pp. 265–353.

[8] G.N. NAIR, R.J. EVANS, I.M.Y. MAREELS AND W. MORAN *Topological feedback entropy and nonlinear stabilization*, IEEE Trans. Autom. Control, 49 (9) (2004), pp. 1585–1597.

[9] C. KAWAN, *Invariance entropy of control systems*. Doctoral thesis, Institut für Mathematik, Universität Augsburg, 2009. submitted.

[10] M. POLLICOTT AND M. YURI, *Dynamical Systems and Ergodic Theory*, Cambridge University Press, 1998.

[11] C. ROBINSON, *Dynamical Systems. Stability, Symbolic Dynamics, and Chaos*, CRC Press, 1999. Second edition.

[12] J. TRUMPF, *On the geometry and parametrization of almost invariant subspaces and observer theory*. Doctoral thesis, Institut für Mathematik, Universität Würzburg, 2002.

[13] J.C. WILLEMS *Almost invariant subspaces: An approach to high gain feedback design—Part I: Almost controlled invariant subspaces*, IEEE Trans. Autom. Control, 26 (1) (1981), pp. 235–252.

[14] M. W. WONHAM, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, 1979.