Systems & Control Letters 20 (1993) 131-139
North-Holland

Null controllability of systems with control constraints and state saturation *

Derong Liu and Anthony N. Michel
Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA

Received 16 March 1992
Revised 30 July 1992

Abstract: Sufficient conditions for the null controllability of discrete-time systems with control constraints and state saturation are presented. This is accomplished in two stages: First, we establish a result for the asymptotic stability of such systems under zero input conditions. Next, we utilize this stability result to obtain a condition for the null controllability of such systems. We include specific examples to demonstrate the applicability of the present results.

Keywords: Controllability; asymptotic stability; nonlinear systems; discrete-time systems; saturation.

1. Introduction

In this paper we investigate the null controllability of systems described by

$$x(k+1) = \text{sat}[Ax(k) + Bu(k)], \quad k = 0, 1, 2, \ldots, \tag{1}$$

where $x(k) \in D^n := [-1, 1]^n$, $u(k) \in D^m = [-1, 1]^m$, $\text{sat}(x) = [\text{sat}(x_1), \ldots, \text{sat}(x_n)]^T$, and

$$\text{sat}(x_i) = \begin{cases} 1, & x_i > 1, \\ x_i, & -1 \leqslant x_i \leqslant 1, \\ -1, & x_i < -1. \end{cases}$$

We refer to such systems as 'linear systems with control constraints and state saturation'. Since we have saturation nonlinearities in (1), it is clear that for any $x(0) \notin D^n$, $x(k) \in D^n$, for $k \geqslant 1$, will always be true. Thus, without loss of generality, we assume that $x(0) \in D^n$. We say that system (1) is null controllable if for any initial state $x(0) \in D^n$ there exists a finite step control sequence $u(k) \in D^m$, $k = 0, 1, 2, \ldots, N$, such that $x(N) = 0$, i.e., the state of system (1) can be steered to the origin from any initial state in D^n by constrained controller $u(k) \in D^m$ in a finite number of steps.

Existing results on the null controllability of linear systems with constrained control are mainly focused on linear systems

$$x(k+1) = Ax(k) + Bu(k), \tag{2}$$

where $A \in R^{n \times n}$, $B \in R^{n \times m}$, $x \in R^n$, $u \in \Omega \subset R^m$ and the control constraint set Ω is convex and compact and contains the origin in its interior. Systems (2) is globally null controllable if its state can be steered to the origin by using $u(k) \in \Omega$, in a finite number of steps, for any given initial state $x(0 \in R^n$. It is shown in [2,7] that system (2) is globally null controllable *if and only if* (i) (A, B) is controllable; and (ii) every eigenvalue λ of A satisfies $|\lambda(A)| \leqslant 1$. The results in [2] and [7] have been extended to linear time-varying systems (see [5,6,9]). In the present paper we will not consider time-varying systems.

Correspondence to: Prof. A.N. Michel, Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA.
* This work was supported by the National Science Foundation under grant ECS 91-07728.

We mention here that there exist systems of type (1), for which rank$[B, AB, \ldots, A^{n-1}B] = n$ and $|\lambda(A)| < 1$ do not imply null controllability in D^n with $u(k) \in D^m$ (see Example 4). This suggests a great difference between the qualitative behavior of (1) and (2), since the necessary and sufficient conditions of null controllability for (2) do not even constitute sufficient conditions for the null controllability of (1). Interestingly, we will also see that it is possible for system (1) to be null controllable even when A is not stable. It is these observations which motivate the current work.

2. Main results

Before presenting our results on the null controllability of system (1), let us consider the system described by (1) under zero-input, given by

$$x(k+1) = \text{sat}[Ax(k)], \quad k = 0, 1, 2, \ldots. \tag{3}$$

We will first establish a stability result for system (3) which is required in the sequel. This result involves the following hypothesis and definition.

Assumption (A1). Let $x_s = \text{sat}(x) = [\text{sat}(x_1), \ldots, \text{sat}(x_n)]^T$ for $x \in R^n$ and let $H \in R^{n \times n}$ denote a symmetric, positive definite matrix. Assume that

$$x_s^T H x_s < x^T H x, \tag{4}$$

whenever $x \notin D^n$, $x \in R^n$.

The next result gives a *necessary and sufficient* condition for matrices to satisfy Assumption (A1). This result is very useful in applications.

Lemma 2.1. *An $n \times n$ positive definite matrix $H = (h_{ij})$ satisfies Assumption* (A1) *if and only if*

$$h_{ii} \geq \sum_{j=1, j \neq i}^{n} |h_{ij}|, \quad i = 1, \ldots, n. \tag{5}$$

Proof. See Appendix.

Definition 2.1. A matrix $W \times R^{n \times n}$ is said to possess *Property P* if it is stable (i.e., all eigenvalues of W are within the unit circle) and if there exists a matrix H satisfying (A1), such that $H - W^T H W$ is positive semidefinite.

We are now in a position to establish the following stability result.

Theorem 2.1. *The equilibrium $x_e = 0$ of system* (3) *is globally asymptotically stable if the matrix A in system* (3) *has Property* P.

Proof. Since matrix A possesses Property P, there exists a matrix H which satisfies (A1), i.e., $[\text{sat}(x)]^T H [\text{sat}(x)] \leq x^T H x$ for any $x \in R^n$, such that $Q := H - A^T H A$ is positive semidefinite. Choose a function $v : R^n \to R$ as $v(x(k)) = x^T(k) H x(k)$ for system (3). This function v is clearly positive definite and radially unbounded. Also, since

$$\begin{aligned} Dv_{(3)}(x(k)) &:= v(x(k+1)) - v(x(k)) \\ &= [\text{sat}(Ax(k))]^T H [\text{sat}(Ax(k))] - x^T(k) H x(k) \leq x^T(k)(A^T H A - H) x(k), \end{aligned}$$

and since $H - A^{\mathrm{T}}HA$ is positive semidefinite, $Dv_{(3)}(x(k))$ is negative semidefinite for all $x(k)$. Therefore, the equilibrium $x_e = 0$ is stable.

To show that the equilibrium $x_e = 0$ of (3) is globally asymptotically stable, we must show that for any $x(0) \in R^n$, $x(k) \to 0$ as $k \to \infty$ (cf. [4]). As mentioned earlier, if $x(0) \notin D^n$, then $x(k) \in D^n$ for all $k \geqslant 1$. Thus we assume without loss of generality that $x(0) \in D^n$.

Let us consider an n consecutive step iteration for the system (3), from $n_0 \geqslant 0$ to $n + n_0$. Without loss of generality, assume that the system (3) saturates at $k = l$, $l \in [n_0, n + n_0)$. In view of (A1), it follows that

$$\begin{aligned} v(x(l+1)) &= x^{\mathrm{T}}(l+1)Hx(l+1) = [\mathrm{sat}(Ax(l))]^{\mathrm{T}}H[\mathrm{sat}(Ax(l))] \\ &< [Ax(l)]^{\mathrm{T}}HAx(l) \leqslant x^{\mathrm{T}}(l)Hx(l) = v(x(l)). \end{aligned}$$

On the other hand, if no saturation occurs during this period, then, using the fact that if $H - A^{\mathrm{T}}HA$ is positive semidefinite, then $H - (A^{\mathrm{T}})^n HA^n$ is positive definite when A is stable (cf. [8]), we thus have

$$\begin{aligned} v(x(n+n_0)) &= x^{\mathrm{T}}(n+n_0)Hx(n+n_0) = [A^n x(n_0)]^{\mathrm{T}} HA^n x(n_0) \\ &= x^{\mathrm{T}}(n_0)(A^{\mathrm{T}})^n HA^n x(n_0) < x^{\mathrm{T}}(n_0)Hx(n_0) = v(x(n_0)). \end{aligned}$$

We can conclude that for the sequence $\{k\colon k = 1, 2, \ldots\}$ there always exists an infinite subsequence $\{k_j\colon j = 1, 2, \ldots\}$, such that $Dv_{(3)}(x(k_j))$ is negative for $x(k_j) \neq 0$, $x(k_j) \in D^n$, and that $v(x(k)) \leqslant v(x(k_j))$ for all $k \geqslant k_j$. Since v is a positive definite quadratic form, it follows that $v(x(k_j)) \to 0$ as $j \to \infty$, and $v(x(k)) \to 0$ as $k \to \infty$. This in turn implies that $x(k) \to 0$ as $k \to \infty$. Thus, the equilibrium $x_e = 0$ of (3) is globally asymptotically stable. □

Note that the result given in Theorem 2.1 can also be found in one of our recent works [3]. We cite it here for the sake of archival completeness.

The following two lemmas are required in the proof of our main result (Theorem 2.2). Their proofs are straightforward and are therefore not included here.

Lemma 2.2. *Suppose $F \in R^{m \times n}$. Then $Fx \in D^m$ for any $x \in D^n$ if and only if $\| F \|_\infty \leqslant 1$, where $\| \cdot \|_\infty$ denotes the matrix norm induced by the l_∞ vector norm.*

Lemma 2.3. *In system* (1), *if (A, B) is controllable, there exists a set $\Psi \subset (D^n)^0$ with $0 \in (\Psi)^0$, such that every initial point in Ψ can be steered to the origin, with its trajectory remaining inside $(D^n)^0$, in a finite number of steps using $u(k) \in D^m$, where $(\Psi)^0$ denotes the interior of Ψ.*

Theorem 2.2. *System* (1) *is null controllable if (A, B) is controllable and if there exists an $F \in R^{m \times n}$, with $\|F\|_\infty \leqslant 1$, such that $G = A + BF$ possesses Property* P.

Proof. From Lemma 2.2, we know that if $\| F \|_\infty \leqslant 1$, then $Fx \in D^m$ for any $x \in D^n$. We can substitute the linear state feedback $u(k) = Fx(k)$ into (1) to obtain

$$x(k+1) = \mathrm{sat}[Ax(k) + BFx(k)] = \mathrm{sat}[Gx(k)], \quad k = 0, 1, 2, \ldots. \tag{6}$$

Since G possesses Property P, we know from Theorem 2.1 that for (6), $x = 0$ is globally asymptotically stable, i.e., $x(k) \to 0$ as $k \to \infty$ for any $x(0) \in D^n$.

The rest of the proof follows directly from Lemma 2.3 and from the fact that on $(D^n)^0$ system (1) is equivalent to system (2). □

Remark 2.1. If the conditions of Theorem 2.2 are satisfied, we have determined a linear state feedback controller which will stabilize the plant and for which the resulting feedback control system will never violate the control constraints and will tend to the origin asymptotically.

Remark 2.2. The sufficient condition for the null controllability of system (1) given in Theorem 2.2 is stated in terms of the feedback law itself. More intrinsic conditions for the null controllability of system (1) are not apparent at this time.

Since system (6) is frequently used as the model of a zero-input fixed-point digital filter with saturation arithmetic, we also have the following results.

Theorem 2.3. *System* (1) *is null controllable if* (A, B) *is controllable and if one of the following conditions holds*:

(i) *there exists an* $F \in R^{m \times n}$, *with* $\| F \|_\infty \leqslant 1$, *such that* $\rho(|G|) < 1$, *where* $\rho(\cdot)$ *denotes spectral radius*, $G = (g_{ij}) = A + BF$, *and* $|G| = (|g_{ij}|)$;

(ii) *there exists an* $F \in R^{m \times n}$, *with* $\| F \|_\infty \leqslant 1$, *such that* $\| G \|_p = \| A + BF \|_p < 1$ *for some* $p \geqslant 1$, *where* $\| \cdot \|_p$ *denotes the matrix norm induced by the* l_p *vector norm.*

Proof. (i) Under the condition that $\rho(|G|) < 1$, system (6) satisfies $x(k) \to 0$ as $k \to \infty$ for any $x(0) \in D^n$. The proof can be found in [1].

(ii) Choose $v(x) = \| x \|_p$. Then v is positive definite and radially unbounded. From the definition of l_p vector norm, we see that

$$v(\text{sat}(x)) = \|\text{sat}(x)\|_p \leqslant \| x \|_p = v(x), \quad \text{for any } x \in R^n.$$

We also have, $v(Gx) = \| Gx \|_p \leqslant \| G \|_p \| x \|_p = v(x)$ for every $x \in R^n$, since $\| G \|_p < 1$. Hence,

$$\begin{aligned} Dv_{(6)}(x(k)) &= v(x(k+1)) - v(x(k)) = v(\text{sat}[Gx(k)]) - v(x(k)) \\ &\leqslant v(Gx(k)) - v(x(k)) < v(x(k)) - v(x(k)) = 0, \end{aligned}$$

i.e., $Dv_{(6)}(x(k))$ is negative definite on R^n. Thus, the equilibrium $x_e = 0$ of system (1) is globally asymptotically stable (cf. [4]).

The rest of the proof follows directly from the proof of Theorem 2.2. □

Remark 2.3. In $G = A + BF = (g_{ij})$, we see that

$$g_{ij} = a_{ij} + \sum_{k=1}^{m} b_{ik} f_{kj}, \quad i = 1, 2, \ldots, n, \; j = 1, 2, \ldots, n.$$

It is clear that $\| A - BF \|_p < 1$ consists of only first order inequalities involving f_{ij} when $p = 1$ or $p = \infty$. Therefore, we can say that $\| A + BF \|_p < 1$ constitutes a 'linear' condition when $p = 1$ or $p = \infty$, since in these cases F can be solved by linear programming [10] when $\| F \|_\infty < 1$. We will give a demonstration of this idea in Section 5, by applying linear programming to a specific second order system (Example 1).

Remark 2.4. Under certain circumstances, the condition $\| G \|_p = \| A + BF \|_p < 1$ in Theorem 2.3(ii) can be relaxed to $\| G \|_p \leqslant 1$. For example, if $p = \infty$, if

$$\max_{i, B_i \neq 0} \left\{ \sum_{j=1}^{n} | g_{ij} | \right\} < 1, \qquad \text{if} \quad \max_{i, B_i = 0} \left\{ \sum_{j=1}^{n} | g_{ij} | \right\} = \max_{i, B_i = 0} \left\{ \sum_{j=1}^{n} | a_{ij} | \right\} \leqslant 1,$$

and if (A, B) is controllable, then system (1) will still be null controllable. This observation allows us to apply the present results to a larger class of systems (e.g., to systems where (A, B) is in the controllable companion form (see Example 3)).

Remark 2.5. The conditions of Theorem 2.2 are less conservative than those in Theorem 2.3 (as can be seen in Example 2 of the next section); however, Theorem 2.3 is much easier to apply than Theorem 2.2.

3. Conclusions and examples

Because of the wide applicability of system (1) to *control systems* and to *fixed-point digital filters*, he null controllability of such systems is of great interest. We emphasize that when the controller is constrained (as in the present case, where $u(k) \in D^m$), system (1) may not be null controllable even when (A, B) is controllable and A is stable.

We solve the problem addressed herein in two steps: (i) we give conditions under which system (1) can be stabilized by constrained controllers $u(k) \in D^m$; and (ii) we prove that if system (1) is stabilizable by constrained controllers and if (A, B) is controllable, then system (1) is null controllable. As pointed out (Remark 2.1), our results (Theorem 2.2 and 2.3) can also serve as conditions for *stabilizability* of system (1), in which case we do not require that (A, B) be controllable.

We demonstrate the applicability of the present results by means of the following examples.

Example 1. In system (1), take

$$A = \begin{pmatrix} 0.4 & 0.5 \\ -0.3 & -1.2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then

$$A + BF = \begin{pmatrix} 0.4 & 0.5 \\ -0.3 + f_1 & -1.2 + f_2 \end{pmatrix}$$

and $\operatorname{rank}[B \ AB] = 2$.

We can see that there exist many $F \in R^{1 \times 2}$ such that $\| F \|_\infty \leqslant 1$, i.e., $| f_1 | + | f_2 | \leqslant 1$, and $\| A + BF \|_\infty < 1$ (refer to the crosshatched region in Figure 1). Therefore, this system is null controllable by Theorem 2.3(ii).

Note that A is unstable, since $\lambda(A) = -1.1, 0.3$.

Example 2. It is easily seen that for system (1) with

$$A = \begin{pmatrix} 0.2 & 1 \\ -1 & 1.8 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{7}$$

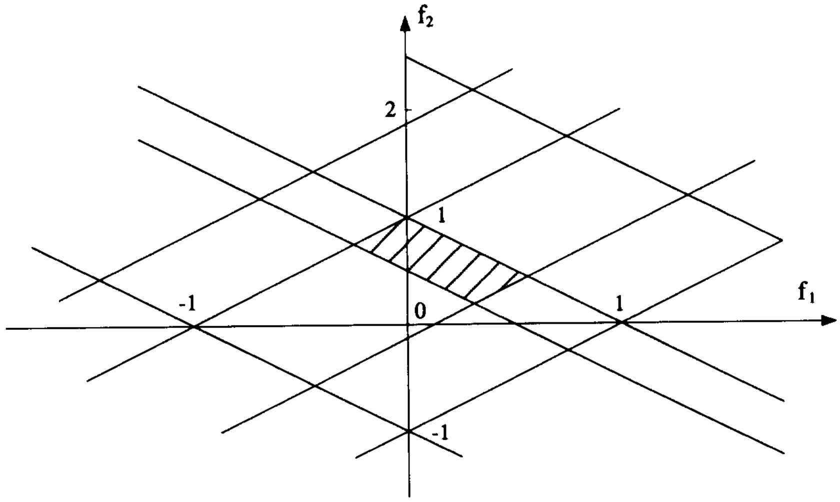

Fig. 1. The existence of the non-void crosshatched set in the $f_1 - f_2$ plane ensures the null controllability of the system in Example 1.

Theorem 2.3 fails. We attempt to apply Theorem 2.2.

Choosing $F = (0.3, -0.7)^T$, we have $\| F \|_\infty = 1$ and

$$G = A + BF = \begin{pmatrix} 0.2 & 1 \\ -0.7 & 1.1 \end{pmatrix}.$$

It can easily be verified that G possesses Property P for this specific choice of F. To see this, we choose matrix H, which satisfies Assumption (A1), as

$$H = \begin{pmatrix} 1 & -0.6 \\ -0.6 & 1.4 \end{pmatrix}.$$

Then,

$$Q = H - A^T HA = \begin{pmatrix} 0.106 & -0.01 \\ -0.01 & 0.026 \end{pmatrix},$$

which is positive definite. Therefore, system (1) with such (A, B) is null controllable.

Example 3. In the present example, we choose a specific second order case for which the domain of null controllability in the parameter plane can be computed *exactly*, by inspection. We then apply Theorems 2.2 and 2.3 to the same example, to determine the quality of the results obtained by these theorems. *We emphasize that, in general, it is not possible to determine the exact domain of null controllability in the parameter plane for the class of nonlinear systems considered herein.*

We consider the system

$$x(k+1) = \text{sat}[Ax(k) + Bu(k)] = \text{sat}\left[\begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix} x(k) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(k)\right]. \tag{8}$$

(a) When $u(k)$ is constrained to $[-1, 1]$, it can be shown by inspection that system (8) is null controllable when (a, b) belongs to the crosshatched region shown in Figure 2.

(b) For conditions $\| A + BF \|_{p=\infty} < 1$ and $\| F \|_\infty \leqslant 1$ to be satisfied, (a, b) must belong to the crosshatched region indicated in Figure 3, which is the region of null controllability for system (8), obtained by Theorem 2.3(ii), when $p = \infty$ (cf. Remark 2.4).

(c) For the conditions of Theorem 2.2 to be satisfied, (a, b) must belong to the crosshatched region shown in Figure 4 (refer to [3] for details).

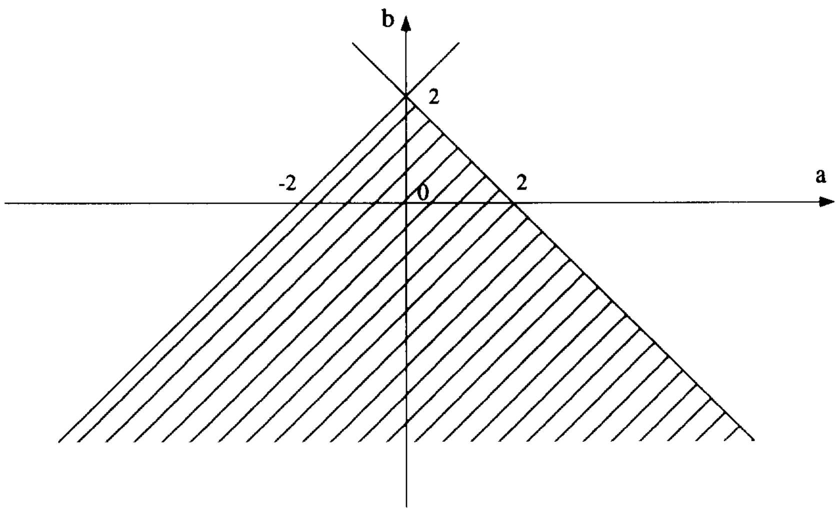

Fig. 2. Domain of null controllability in the parameter plane obtained by inspection (Example 3).

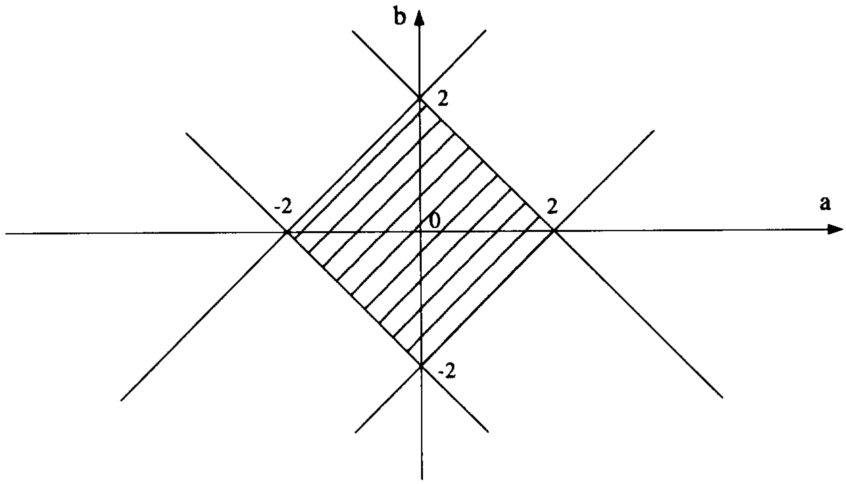

Fig. 3. Domain of null controllability in the parameter plane obtained by Theorem 2.3(ii) when $p=\infty$ (Example 3).

Note that the crosshatched region in Figure 4 contains the crosshatched region of Figure 3 and is contained in the crosshatched region of Figure 2.

Example 4. In the present example, we demonstrate the validity of our original claim that for system (1) the conditions that A be stable and $\text{rank}[B, AB, \ldots, A^{n-1}B]=n$ do not imply null controllability. Specifically, for the system

$$x(k+1)=\text{sat}[Ax(k)+Bu(k)]=\text{sat}\left[\begin{pmatrix} 1.6 & 0.2 \\ -30 & -3.3 \end{pmatrix}x(k)+\begin{pmatrix} 0 \\ 1 \end{pmatrix}u(k)\right],$$

where $x(0)\in D^2=[-1, 1]^2$ and $u(k)\in D^1=[-1, 1]$, we have $\text{rank}[B\ AB]=2$ and $\lambda_1(A)=-0.8$ and $\lambda_2(A)=-0.9$. However, this system is not null controllable, since when we choose $x(0)=(a, b)^{\text{T}}$ or $(-a, b)^{\text{T}}$, with $0.75\leqslant a\leqslant 1$ and $-1\leqslant b\leqslant 1$, the state of this system will remain at $(1, -1)^{\text{T}}$ or $(-1, 1)^{\text{T}}$, respectively, for any $u\in D^1$.

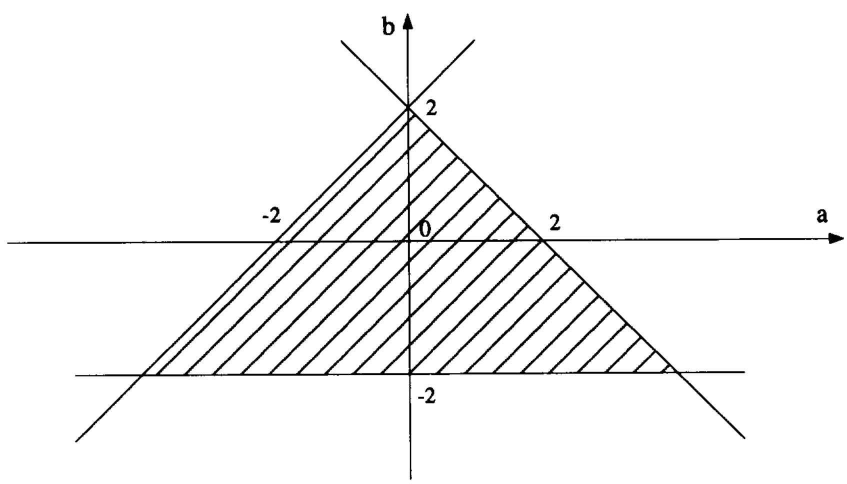

Fig. 4. Domain of null controllability in the parameter plane obtained by Theorem 2.2 (Example 3).

Appendix

Proof of Lemma 2.1. We first introduce the following notation. Denote

$$x_s = [\text{sat}(x_1), \ldots, \text{sat}(x_n)]^T = Ex,$$

where $E = \text{diag}(e_1, e_2, \ldots, e_n)$, $e_i = 1$ when $|x_i| \leqslant 1$, and $e_i = 1/|x_i|$ when $|x_i| > 1$.

Then, we have $x^T Hx - x_s^T Hx_s = x^T(H - EHE)x$.

Sufficiency: Suppose $x = (x_1, x_2, \ldots, x_n)^T$, $|x_k| > 1$ and $|x_i| \leqslant 1$ for $i \neq k$ $(x \notin D^n)$. We have $0 < e_k < 1$, and $e_i = 1$ for $i \neq k$. Therefore,

$$H - EHE = \begin{bmatrix} 0 & \cdots & 0 & h_{1k}(1-e_k) & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & h_{k-1,k}(1-e_k) & 0 & \cdots & 0 \\ h_{k1}(1-e_k) & \cdots & h_{k,k-1}(1-e_k) & h_{kk}(1-e_k^2) & h_{k,k+1}(1-e_k) & \cdots & h_{kn}(1-e_k) \\ 0 & \cdots & 0 & h_{k+1,k}(1-e_k) & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & h_{nk}(1-e_k) & 0 & \cdots & 0 \end{bmatrix},$$

and

$$x^T(H - EHE)x = (1 - e_k)\left(h_{kk}(1 + e_k)x_k^2 + 2\sum_{i=1, i\neq k}^{n} h_{ik}x_i x_k\right). \tag{9}$$

Note that in the above equation we have used the fact that $h_{ij} = h_{ji}$. From $|x_i| \leqslant 1$ for $i \neq k$, $|x_k| > 1$, and $e_k|x_k| = 1$, we have $2|x_i x_k| < (1 + |x_k|)|x_k| = (1 + e_k)x_k^2$. Hence, from (9), we have

$$x^T(H - EHE)x \geqslant (1 - e_k)\left(h_{kk}(1 + e_k)x_k^2 - 2\sum_{i=1, i\neq k}^{n} |h_{ik}x_i x_k|\right)$$
$$> (1 - e_k^2)x_k^2\left(h_{kk} - \sum_{i=1, i\neq k}^{n} |h_{ik}|\right) \geqslant 0,$$

i.e., $x^T Hx > x^T EHEx = x_s^T Hx_s$.

Denote $M = \{1, 2, \ldots, m\}$ for *any* m, $0 < m \leqslant n$, and $N = \{k_i\colon 0 < k_i \leqslant n,\ k_i \neq k_j$, when $i \neq j,\ i \in M\}$. Now suppose that $x = (x_1, x_2, \ldots, x_n)^T$, $|x_k| > 1$ for $k \in N$ and $|x_i| \leqslant 1$ for $i \notin N$ $(x \notin D^n)$. Following the same procedure as above, we have

$$\begin{aligned} x^T(H - EHE)x \\ &= \sum_{k\in N}(1 - e_k)\left(h_{kk}(1 + e_k)x_k^2 + 2\sum_{i=1, i\notin N}^{n} h_{ik}x_i x_k\right) + \sum_{k\in N}\sum_{l\in N, l\neq k} h_{kl}x_k x_l(1 - e_k e_l) \\ &\geqslant \sum_{k\in N}(1 - e_k)\left(h_{kk}(1 + e_k)x_k^2 - 2\sum_{i=1, i\notin N}^{n} |h_{ik}x_i x_k|\right) + \sum_{k\in N}\sum_{l\in N, l\neq k} h_{kl}x_k x_l(1 - e_k e_l) \\ &> \sum_{k\in N}(1 - e_k^2)x_k^2\left(h_{kk} - \sum_{i=1, i\notin N}^{n} |h_{ik}|\right) + \sum_{k\in N}\sum_{l\in N, l\neq k} h_{kl}x_k x_l(1 - e_k e_l) \\ &= \sum_{k\in N}(1 - e_k^2)x_k^2\left(h_{kk} - \sum_{i=1, i\neq k}^{n} |h_{ik}|\right) + \sum_{k\in N}(1 - e_k^2)x_k^2 \sum_{i\in N, i\neq k} |h_{ik}| \\ &\quad + \sum_{k\in N}\sum_{l\in N, l\neq k} h_{kl}x_k x_l(1 - e_k e_l). \end{aligned} \tag{10}$$

The first summation of the right hand side in (10) is nonnegative by assumption. Considering the last two terms in the right hand side of (10), by noting that $0 < e_k < 1$ and $e_k |x_k| = 1$ for $k \in N$, we have

$$\begin{aligned}
&\sum_{k \in N} (1 - e_k^2) x_k^2 \sum_{i \in N, i \neq k} |h_{ik}| + \sum_{k \in N} \sum_{l \in N, l \neq k} h_{kl} x_k x_l (1 - e_k e_l) \\
&\quad \geqslant \sum_{k \in N} \sum_{l \in N, l \neq k} (1 - e_k^2) x_k^2 |h_{kl}| - \sum_{k \in N} \sum_{l \in N, l \neq k} |h_{kl} x_k x_l| (1 - e_k e_l) \\
&\quad = \sum_{k \in N} \sum_{l \in N, l \neq k} |h_{kl} x_k| (|x_k| - e_k - |x_l| + e_k) = \sum_{k \in N} \sum_{l \in N, l \neq k} |h_{kl}| x_k^2 - \sum_{k \in N} \sum_{l \in N, l \neq k} |h_{kl} x_k x_l| \\
&\quad = \sum_{k \in N} \sum_{l \in N, l > k} |h_{kl}| (x_k^2 + x_l^2) - 2 \sum_{k \in N} \sum_{l \in N, l > k} |h_{kl} x_k x_l| = \sum_{k \in N} \sum_{l \in N, l > k} |h_{kl}| (|x_k| - |x_l|)^2 \geqslant 0.
\end{aligned}$$

Therefore, $x^{\mathrm{T}} H x - x_{\mathrm{s}}^{\mathrm{T}} H x_{\mathrm{s}} = x^{\mathrm{T}} (H - EHE) x > 0$, for any $x \in R^n$ such that $x \notin D^n$. This proves the sufficiency.

Necessity: It suffices to show that if (5) does not hold, there always exist some points $x \notin D^n$, such that $x^{\mathrm{T}} H x \leqslant x_{\mathrm{s}}^{\mathrm{T}} H x_{\mathrm{s}}$. Suppose that (5) does not hold for $i = k$, i.e.,

$$\delta := \sum_{j=1, j \neq k}^{n} |h_{kj}| - h_{kk} > 0.$$

Let us choose $|x_k| = 1 + \xi$, $\xi > 0$, and $x_i = -\operatorname{sign}(h_{ik} x_k)$, $i \neq k$. Then, $x = (x_1, \ldots, x_n)^{\mathrm{T}} \notin D^n$ and (9) becomes

$$\begin{aligned}
x^T (H - EHE) x &= (1 - e_k) \left(h_{kk} (1 + e_k) x_k^2 - 2 \sum_{i=1, i \neq k}^{n} |h_{ik} x_k| \right) \\
&= (1 - e_k) |x_k| \left(h_{kk} \xi + 2 h_{kk} - 2 \sum_{i=1, i \neq k}^{n} |h_{ki}| \right) = (1 - e_k) |x_k| (h_{kk} \xi - 2\delta).
\end{aligned}$$

Clearly, when we choose $0 < \xi \leqslant 2\delta / h_{kk}$, we have $x^{\mathrm{T}} H x - x_{\mathrm{s}}^{\mathrm{T}} H x_{\mathrm{s}} = x^{\mathrm{T}} (H - EHE) x \leqslant 0$. Note here that $h_{kk} > 0$ since H is positive definite. This proves the necessity. □

References

[1] P.H. Bauer and E.I. Jury, A stability analysis of two-dimensional nonlinear digital state-space filters, *IEEE Trans. Acoust. Speech Signal Process.* **38** (1990) 1578–1586.

[2] E.B. Lee and L. Markus, *Foundations of Optimal Control Theory* (John Wiley and Sons, New York, 1967).

[3] Derong Liu and A.N. Michel, Asymptotic stability of discrete-time systems with saturation nonlinearities, submitted to the *31st IEEE Conference on Decision and Control*.

[4] R.K. Miller and A.N. Michel, *Ordinary Differential Equations* (Academic Press, New York, 1982).

[5] L. Pandolfi, Linear control systems: Controllability with constrained controls, *J. Optim. Theory Appl.* **19** (1976) 577–585.

[6] W.E. Schmitendorf and B.R. Barmish, Null controllability of linear systems with constrained controls, *SIAM J. Control Optim.* **18** (1980) 327–345.

[7] E.D. Sontag, An algebraic approach to bounded controllability of linear systems, *Internat. J. Control* **39** (1984) 181–188.

[8] P.P. Vaidyanathan and V. Liu, An improved sufficient condition for absence of limit cycles in digital filters, *IEEE Trans. Circuits and Systems* **34** (1987) 319–322.

[9] R.P. Van Til and W.E. Schmitendorf, Constrained controllability of discrete-time systems, *Internat. J. Control* **43** (1986) 943–956.

[10] G.R. Walsh, *An Introduction to Linear Programming* (John Wiley and Sons, New York, 1985).