Working Paper Series

Controllability of Discrete–Time Two dimensional Single Input Recurrent Neural Networks

Thomas Steinberger
Lucas Zinner

Working Paper No. 56
December 1999

Working Paper Series

December 1999

SFB
‘Adaptive Information Systems and Modelling in Economics and Management Science’

Vienna University of Economics
and Business Administration
Augasse 2–6, 1090 Wien, Austria

in cooperation with
University of Vienna
Vienna University of Technology

http://www.wu-wien.ac.at/am

This piece of research was supported by the Austrian Science Foundation (FWF) under grant SFB#010 (‘Adaptive Information Systems and Modelling in Economics and Management Science’).

CONTROLLABILITY OF DISCRETE–TIME TWO DIMENSIONAL SINGLE INPUT RECURRENT NEURAL NETWORKS

THOMAS STEINBERGER AND LUCAS ZINNER

ABSTRACT. This paper presents a complete characterization of controllability and reachability for the class of discrete–time recurrent neural networks with state space dimension 2 and single input.

1. INTRODUCTION

In this paper we deal with control systems in discrete time. In general, by an n–dimensional, m–input recurrent σ–net we mean a discrete time control system of the form

$$\boldsymbol{x}(t+1) = \vec{\sigma}^{(n)}(A\boldsymbol{x}(t) + Bu(t)), \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ and the map $\sigma : \mathbb{R} \to \mathbb{R}$ denotes a sigmoid function. Here for each positive integer n we use $\vec{\sigma}^{(n)}$ to denote the diagonal mapping

$$\vec{\sigma}^{(n)} : \mathbb{R}^n \to \mathbb{R}^n : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \sigma(x_1) \\ \vdots \\ \sigma(x_n) \end{pmatrix}$$

The spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ are called respectively the state space and the input–value space of the net. We mention that linear systems studied in control theory are the σ–nets for which σ is the identity function.

In neural networks theory one interprets the equation (1) as representing the evolution of an ensemble of n "neurons", where each coordinate x_i of x is a real–valued variable which represent the internal state of the ith neuron, and each u_i, $i = 1, \ldots, m$ of u is the external input signal. The coefficients A_{ij}, B_{ij} denote the weights or "synaptic strengths"of the various connections. Finally the transformation $\sigma : \mathbb{R} \to \mathbb{R}$ is usually called the "activation function". One example of such functions is the hyperbolic tangent $\sigma = \tanh$.

Questions of parameter identifiability and observability were studied for such models in [3] and [4]. Here we focus on problems of controlability. One of

---

Key words and phrases. Controllability, Discrete Time Recurrent Neural Networks .

Research of the authors was supported by *FWF*–grants SFB10 and P 11213-Mat of the Austrian Ministry of Science.

the first contribution in this area was a paper of F. Albertini and P. Dai Pra, cf. [2], where they dealt with the study of the forward accessibility. Recall that a system (1) is called *forward accessible* if from each initial state it is possible to reach by using appropriate inputs $u(\cdot)$ an open set of the state space. They showed that forward accessibility holds provided that a certain independence property holds for σ and the matrix B belongs to a certain class $\boldsymbol{B}_{n,m}$ of matrices which were introduced in [1], see also [3].
In [7] (see also [5]) E. Sontag and H. Sussmann studied the question of controllability in the continuous–time setting. A system is called *completely controllable* if from each initial state it is possible to reach, by using appropriate inputs $u(\cdot)$, the entire state space, not only some — maybe very "small" — open set. In the continuous–time setting the equation (1) may be read as

$$\dot{\boldsymbol{x}}(t) = \vec{\sigma}^{(n)}(A\boldsymbol{x}(t) + Bu(t)),$$

where A and B are as above and σ has to be at least locally Lipschitz. The main step in their proof is that they could show that the convex hull of the set $\{\vec{\sigma}^{(n)}(\vec{a} + Bu), u \in \mathbb{R}^m\}$ contains an open (in $\mathbb{R}^n$) neighborhood of zero. This holds if $B \in \boldsymbol{B}_{n,m}$, A arbitrary and $\sigma = \tanh$, but fails if tanh is replaced by say arctan.
In the linear case, that is if σ denotes the identity function, the system is controllable if and only if $\operatorname{rank} \boldsymbol{R}(A,B) = n$, where

$$\boldsymbol{R}(A,B) := [B, AB, A^2B, \ldots, A^{n-1}B].$$

For a proof we refer the reader to [6].
In contrast to the above mentioned results we showed in [8] that in the discrete–time case the assumptions for complete controllability are quite restrictive, more precisely we showed that a system is completely controllable if and only if the matrix B has full rank n. In this paper we focus on two dimensional recurrent nets with single input, that is

$$\vec{\sigma}^{(2)} : \mathbb{R}^2 \to \mathbb{R}^2 : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \vec{\sigma}^{(2)}\left(A\begin{pmatrix} x \\ y \end{pmatrix} + Bu\right),$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

and $u \in \mathbb{R}$. We call B admissible if $b_i \neq 0$ for $i = 1, 2$. Note that this is more general than to assume $B \in \boldsymbol{B}_{2,1}$.

2. Notations and Basic Definitions

To begin with let us recall the basic definitions we use.

Definition 2.1. Let $t = 0, 1, 2, \ldots$. A system Σ given by

$$\boldsymbol{x}(t+1) = \vec{\sigma}^{(n)}(A\boldsymbol{x}(t) + Bu(t))$$

is called recurrent neural network (RNN) with initial state $\boldsymbol{x}_0$ if $\boldsymbol{x}(0) = \boldsymbol{x}_0$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $\sigma : \mathbb{R} \to \mathbb{R}$ is a sigmoid function, that is σ is odd, absolutely bounded by 1 and strictly increasing. Furthermore $\mathcal{X} = (-1,1)^n$ denotes the state space, where $(-1,1)^n$ is defined to be n times of the cartesian product of the intervall $(-1,1)$.

We will focus on the case where $n = 2$ and $m = 1$ in the definition above and write $\vec{\sigma}$ instead of $\vec{\sigma}^{(2)}$. Let us note that all the usually used functions such as tanh and $\frac{2}{\pi}$ arctan fulfill the conditions cited above. In fact it is sufficient that the function σ is assumed to be bounded by some constant. In the following we will use either Greek or bold letters to indicate elements of $\mathcal{X}$ or $\mathbb{R}^2$.

Definition 2.2. Let ζ, η be in $\mathcal{X}$. Then ζ can be reached from η iff there exist $t \in \mathbb{N}$ such that $\boldsymbol{x}(0) = \eta$ and $\boldsymbol{x}(t) = \zeta$ with appropriate choosen $u(0), \ldots, u(t-1)$. One says ζ can be reached from η (or η can be controlled to ζ) if this happens for at least one t.

We use the notation $\eta \rightsquigarrow \zeta$ to indicate that the state ζ can be reached from η. Sometimes we write $\eta \overset{T}{\rightsquigarrow} \zeta$ if ζ can be reached from η in time T. We denote

$$\mathcal{C}^t(\zeta) := \{\eta \in \mathcal{X} : \eta \overset{s}{\rightsquigarrow} \zeta \text{ with } s \leq t\} \quad \text{and} \quad \mathcal{C}(\zeta) = \bigcup_{t=0}^{\infty} \mathcal{C}^t(\zeta)$$

and

$$\mathcal{R}^t(\eta) := \{\zeta \in \mathcal{X} : \eta \overset{s}{\rightsquigarrow} \zeta \text{ with } s \leq t\} \quad \text{and} \quad \mathcal{R}(\eta) = \bigcup_{t=0}^{\infty} \mathcal{R}^t(\eta).$$

Finally denote

$$\mathcal{R} = \bigcup_{\eta \in \mathcal{X}} \mathcal{R}(\eta).$$

Remark 2.1. It is very easy to see that $\eta \rightsquigarrow \zeta$ and $\zeta \rightsquigarrow \xi$ implies $\eta \rightsquigarrow \xi$. It is also clear that if $\eta \overset{T}{\rightsquigarrow} \zeta$ for some $T > 0$ and if $0 < t < T$, then there is some $\xi \in \mathcal{X}$ such that $\eta \overset{t}{\rightsquigarrow} \xi$ and $\xi \overset{T-t}{\rightsquigarrow} \zeta$.
But it is not necessarily true that $\eta \rightsquigarrow \zeta$ implies $\zeta \rightsquigarrow \eta$.

Definition 2.3. A system Σ is said to be completely controllable if for each $\eta, \zeta \in \mathcal{X}$ it holds that $\eta \rightsquigarrow \zeta$. It is said to be completely controllable in time T if for each $\eta, \zeta \in \mathcal{X}$ it holds that $\eta \in \mathcal{C}^T(\zeta)$ and vice versa.

In [8] the following is proved.

Theorem 2.1. *Let Σ be a RNN. Then Σ is completely controllable if and only if the $n \times m$–matrix B has full rank n independent of the matrix A. Especially if $n = 2$ and $m = 1$ then $\mathcal{R} \subsetneq \mathcal{X}$.*

On the other hand we will show in the next section that $\eta \rightsquigarrow 0$ for all $\eta \in \mathcal{X}$ provided A and B fulfill certain properties.

3. Statement and Proof of the Main Results

The following proposition shows that $\mathbf{0}$ can be reach from arbitrary $\boldsymbol{x} \in \mathcal{X}$ in time 2 if appropriate rank conditions are supposed.

Proposition 3.1. *Let Σ be a two dimensional single input recurrent neural network, let B be admissible. If* $\operatorname{rank}[AB, B] = 2$ *then* $\mathcal{C}^2(\mathbf{0}) = \mathcal{X}$.

Proof. Let
$$\boldsymbol{x}_0 = \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$
Then
$$\boldsymbol{x}(1) = \begin{pmatrix} \sigma(a_{11}x + a_{12}y + b_1 u(0)) \\ \sigma(a_{21}x + a_{22}y + b_2 u(0)) \end{pmatrix}$$
and
$$\boldsymbol{x}(2) = \begin{pmatrix} \sigma(a_{11}\sigma(a_{11}x + a_{12}y + b_1u(0)) + a_{12}\sigma(a_{21}x + a_{22}y + b_2u(0)) + b_1u(1)) \\ \sigma(a_{21}\sigma(a_{11}x + a_{12}y + b_1u(0)) + a_{22}\sigma(a_{21}x + a_{22}y + b_2u(0)) + b_2u(1)) \end{pmatrix}$$
We have to show that $\boldsymbol{x}(2) = \mathbf{0}$ for appropriate choosen $u(0), u(1)$.
Case i: Assume $b_1 \neq \pm b_2$.
Then
$$\gamma : \mathbb{R} \to \mathcal{X}, u \mapsto \begin{pmatrix} \sigma(a_{11}x + a_{12}y + b_1u) \\ \sigma(a_{21}x + a_{22}y + b_2u) \end{pmatrix}$$
defines a curve which crosses each line through $\mathbf{0}$ at least one time. If $x_0 \in \ker A$ there is nothing to show since we can choose $u(0) = 0$ to gain our result. Though let us assume $\boldsymbol{x}_0 \notin \ker A$ and $\mathbf{0} \notin \gamma^*$, where γ^* denotes the graph of γ as a subset of $\mathcal{X}$. Since γ crosses each line through $\mathbf{0}$, for each vector $\boldsymbol{v} \in \mathbb{R}^2$ we can find $u(0)$ such that $\gamma(u(0)) = \lambda \cdot \boldsymbol{v}$ for some λ. Now choose $\boldsymbol{v}$ such that $A\boldsymbol{v} = B$ we find some $u(0)$ and get
$$x(2) = \begin{pmatrix} \sigma(\lambda b_1 + b_1u(1)) \\ \sigma(\lambda b_2 + b_2u(1)) \end{pmatrix}.$$
Note that this can be done, since $B \notin \ker A$, otherwise $\operatorname{rank}[AB, B] < 2$. Taking $u(1) = -\lambda$ yields to the desired result.
Case ii: Assume $b_1 = \pm b_2$.
We prove the result for $b_1 = b_2$. If $b_1 = -b_2$ the proof is similar. Set $b_i = b$. We write
$$\boldsymbol{x}(1) = \begin{pmatrix} \sigma(a_{11}x + a_{12}y + bu(0)) \\ \sigma(a_{21}x + a_{22}y + bu(0)) \end{pmatrix}.$$
If $a_{11}x + a_{12}y = a_{21}x + a_{22}y$ we may take $u(0) = -\frac{1}{b}(a_{21}x + a_{22}y)$ so that $\boldsymbol{x}(1) = \mathbf{0}$ and we are done.
Though let us asssume $a_{11}x + a_{12}y \neq a_{21}x + a_{22}y$. Then
$$\gamma : \mathbb{R} \to \mathcal{X}, u \mapsto \begin{pmatrix} \sigma(a_{11}x + a_{12}y + bu) \\ \sigma(a_{21}x + a_{22}y + bu) \end{pmatrix}$$

gives a curve which crosses each line through $\mathbf{0}$ at least one time except the diagonal $\Delta = \{\boldsymbol{v} = (v_1, v_2) \in \mathcal{X} : v_1 = v_2\}$. Since $B \notin \ker A$ there exists a $\boldsymbol{v}$ such that $A\boldsymbol{v} = B$ and $\boldsymbol{v} \notin \Delta$. Again we find $u(0)$ such that $\gamma(u(0)) = \lambda \cdot \boldsymbol{v}$. Choosing $u(1) = -\lambda$ gives $\boldsymbol{x}(2) = \mathbf{0}$ as desired. $\square$

Example 3.1. Let

$$\boldsymbol{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

If $x_0 \neq y_0$, then

$$\boldsymbol{x}(t+1) = \begin{pmatrix} \sigma(x(t) + u(t)) \\ \sigma(y(t) + u(t)) \end{pmatrix} \notin \Delta \quad \text{for all } t = 0, 1, \ldots.$$

Though obviously $\mathcal{C}(\mathbf{0}) = \Delta \subsetneq \mathcal{X}$. In fact if $x_0 > y_0$ the region $\mathcal{R}(\boldsymbol{x}_0)$ is contained in $\{\boldsymbol{x} = (x, y) \in \mathcal{X} : x > y\}$. Note that in this example the rank condition is not fulfilled, since $\operatorname{rank}[AB, B] = 1$.

Example 3.2. Let

$$\boldsymbol{x}_0 = \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

such that $\sigma(x - 2y) > 1/2$. Then

$$\boldsymbol{x}(1) = \begin{pmatrix} \sigma(x + u(0)) \\ \sigma(x - 2y) \end{pmatrix}$$

and

$$\boldsymbol{x}(2) = \begin{pmatrix} \sigma(\sigma(x + u(0)) + u(1)) \\ \sigma(\sigma(x + u(0)) - 2\sigma(x - 2y)) \end{pmatrix}.$$

If $\boldsymbol{x}(2) = \mathbf{0}$ we would have $\sigma(x + u(0)) = 2\sigma(x - 2y) > 1$ by assumption on x and y. But σ is bounded by 1 which yields a contradiction. Note that in this case the rank condition is fulfilled, since $\operatorname{rank}[AB, B] = 2$, but B is not admissible.

The next proposition shows that skipping the rank condition we can not hope to get controllability of $\mathbf{0}$ by increasing time.

Proposition 3.2. *Let Σ be a two dimensional single input recurrent neural network, let B be admissible. If $\boldsymbol{x} \overset{2}{\not\rightsquigarrow} \mathbf{0}$ then $\boldsymbol{x} \overset{t}{\not\rightsquigarrow} \mathbf{0}$ for all $t < \infty$.*

Proof. Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and we suppose there exists some $t > 2$ such that

$$\boldsymbol{x} \overset{t}{\rightsquigarrow} \mathbf{0}, \quad \text{but} \quad \boldsymbol{x} \overset{t-1}{\not\rightsquigarrow} \mathbf{0}.$$

Since $\boldsymbol{x} \overset{2}{\not\rightsquigarrow} \mathbf{0}$ it follows by Proposition 3.1 that $\operatorname{rank}[AB, B] \leq 1$.
Case i: Assume $AB = 0$, that is $B \in \ker A$.
Suppose there exists t such that $\boldsymbol{x}(t+1) = \mathbf{0}$ we have $A\boldsymbol{x}(t) = -Bu(t)$ where $\boldsymbol{x}(t)$ is the state gained after t steps and $u(t) \in \mathbb{R}$ is the appropriate choosen

input. It follows that $B \in \operatorname{image} A$ and hence $B \in \operatorname{image} A \cap \ker A$. This yields to a contradiction to the admissability assumption on B.

Case ii: Assume $AB = \lambda B$ for some $\lambda \in \mathbb{R}$.
If $B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ and $b_1 \neq \pm b_2$ we can follow the proof of Proposition 3.1 to get $\zeta \overset{2}{\rightsquigarrow} \mathbf{0}$ for all $\zeta \in \mathcal{X}$.
Therefore we assume $b_1 = \pm b_2$. Again we only prove the case $b_1 = b_2$ and write $b_i = b$. Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Then $\operatorname{rank}[AB, B] \leq 1$ implies $a_{11} - a_{21} = a_{22} - a_{12}$. If $a_{11} = a_{21}$ we may choose $u(0) = -\frac{1}{b}(a_{11}x + a_{12}y)$ and we are done.
Suppose $a_{11} \neq a_{21}$. By assumption there exists some t such that

$$\mathbf{0} = \boldsymbol{x}(t+1) = \begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} \sigma(a_{11}x(t) + a_{12}y(t) + bu(t)) \\ \sigma(a_{21}x(t) + a_{22}y(t) + b_2u(t)) \end{pmatrix}.$$

This is equivalent to $(a_{11} - a_{21})x(t) = (a_{22} - a_{12})y(t)$ and since $a_{11} \neq a_{21}$ we get $x(t) = y(t)$. Applying the same argument we conclude that $x(t-1) = y(t-1)$. Though choosing $u(t-1) = -\frac{1}{b}(a_{11}x(t-1) + a_{12}y(t-1))$ we get $\boldsymbol{x}(t) = \mathbf{0}$, though $\boldsymbol{x} \overset{t-1}{\rightsquigarrow} \mathbf{0}$ and we are done. ☐

The next indicates that some rank condition is necessary if one is interested in "large" regions of reachability.

Proposition 3.3. *Let Σ be a two dimensional single input recurrent neural network, let B be admissible. If $\operatorname{rank}[A, B] < 2$ then $\mathcal{R} = \mathcal{R}(\boldsymbol{x}) = \mathcal{R}^1(\boldsymbol{x})$ for all $\boldsymbol{x}$ and $\mathcal{R}^1(\boldsymbol{x})$ equals γ^* where γ denotes the curve*

$$\gamma : \mathbb{R} \to \mathcal{X}, \quad u \mapsto \begin{pmatrix} \sigma(b_1u) \\ \sigma(b_2u) \end{pmatrix}.$$

Proof. Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Since $\operatorname{rank}[A, B] < 2$ it follows that there exist λ and μ such that

$$\lambda \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = \mu \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Hence for all $\boldsymbol{v} \in \mathbb{R}^2$ we find $u \in \mathbb{R}$ such that $A\boldsymbol{v} = Bu$ which proves the proposition. ☐

Remark 3.1. We note that the conclusion of Proposition 3.3 does not hold if $\operatorname{rank}[A, B] < 2$ is replaced by $\operatorname{rank}[AB, B] < 2$. Example 3.1 shows that $\mathcal{R}(\boldsymbol{x}) \neq \mathcal{R}(\mathbf{0})$ provided $\boldsymbol{x} \notin \Delta$. In fact $\operatorname{rank}[A, B] < 2$ implies $\operatorname{rank}[AB, B] < 2$, but the converse fails.

Lemma 3.1. *Let Σ be a two dimensional single input recurrent neural network, let B be admissible. Let $\boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X}$. We define the curve*

$$\gamma_{\boldsymbol{x}} : \mathbb{R} \to \mathcal{X}, \quad u \mapsto \vec{\sigma}(A\boldsymbol{x} + Bu).$$

If there exist u_1 and u_2 such that $\gamma_{\boldsymbol{x}_1}(u_1) = \gamma_{\boldsymbol{x}_2}(u_2)$ then $\gamma^*_{\boldsymbol{x}_1} = \gamma^*_{\boldsymbol{x}_2}$.

Proof. If $\gamma_{\boldsymbol{x}_1}(u_1) = \gamma_{\boldsymbol{x}_2}(u_2)$ it follows that $A\boldsymbol{x}_1 + Bu_1 = A\boldsymbol{x}_2 + Bu_2$ and hence $A\boldsymbol{x}_1 = A\boldsymbol{x}_2 + B(u_2 - u_1)$. Therefore

$$\gamma_{\boldsymbol{x}_1}(u) = \vec{\sigma}(A\boldsymbol{x}_1 + Bu) = \vec{\sigma}(A\boldsymbol{x}_2 + B(u_2 - u_1 + u)) = \gamma_{\boldsymbol{x}_2}(u_2 - u_1 + u)$$

which gives the desired result. □

Proposition 3.4. *Let Σ be a two dimensional single input recurrent neural network, let B be admissible. Let $\operatorname{rank} A < 2$ and $\operatorname{rank}[A, B] = 2$. Then for all $\boldsymbol{x} \in \mathcal{X}$ we get $\mathcal{R}(\boldsymbol{x}) = \mathcal{R}^2(\boldsymbol{x})$, that means if any state can be reached from $\boldsymbol{x}$ in finite time it can be reach in at most two steps of iteration.*

Proof. Let $\zeta \in \mathcal{R}(\boldsymbol{x})$, then there exist $t < \infty$ and $u(0), \ldots, u(t)$ such that $\boldsymbol{x}(t+1) = \zeta$ or equivalent $\zeta \in \gamma^*$ where $\gamma : u \mapsto \vec{\sigma}(A\boldsymbol{x}(t) + Bu)$.
Claim 1: There exist u' and v' such that $\gamma(u') = c_0(v')$ where

$$c_0 : \mathbb{R} \to \mathcal{X}, \quad v \mapsto \vec{\sigma}\big(A\vec{\sigma}(A\boldsymbol{x}_0 + Bv)\big).$$

The curve c_0 may be interpreted as the curve of starting points with respect to $\boldsymbol{x}(2)$.
Suppose this claim is allready proved we get

$$\vec{\sigma}\big(A\vec{\sigma}(A\boldsymbol{x}_0 + Bv')\big) = \vec{\sigma}(A\boldsymbol{x}(t) + Bu')$$

or equivalent

$$A\boldsymbol{x}(t) = A\vec{\sigma}(A\boldsymbol{x}_0 + Bv') - Bu'$$

Since $\zeta \in \gamma^*$ there is an input $u(t)$ such that $\zeta = \vec{\sigma}(Ax(t) + Bu(t))$. Now choose new inputs $\tilde{u}(0) = v'$ and $\tilde{u}(1) = u(t) - u'$ we get

$$\begin{aligned}\tilde{\boldsymbol{x}}(2) &= \vec{\sigma}(A\tilde{\boldsymbol{x}}(1) + B\tilde{u}(1)) = \vec{\sigma}\big(A\vec{\sigma}(A\boldsymbol{x}_0 + B\tilde{u}(0)) + B\tilde{u}(1)\big)\\ &= \vec{\sigma}\big(A\boldsymbol{x}(t) + Bu' + B\tilde{u}(1)\big) = \vec{\sigma}\big(A\boldsymbol{x}(t) + B\tilde{u}(t)\big) = \boldsymbol{x}(t+1) = \zeta\end{aligned}$$

and the propositon follows.
To prove claim 1 we will first prove the following
Claim 2: γ crosses the curve c_{t-1} exactly one time, where

$$c_{t-1} : \mathbb{R} \to \mathcal{X}, \quad v \mapsto \vec{\sigma}\big(A\vec{\sigma}(A\boldsymbol{x}(t-1) + Bv)\big),$$

that is, suppose there exist u_1, u_2 and v_1, v_2 such that $\gamma(u_1) = c_{t-1}(v_1)$ and $\gamma(u_2) = c_{t-1}(v_2)$ then $u_1 = u_2$ and $v_1 = v_2$.
It is clear by the definition that γ crosses the curve c_{t-1} since $\gamma(0) = c_{t-1}(u(t-1))$. Now suppose there exist $u \neq 0$ and $v \neq u(t-1)$ such that $\gamma(u) = c_{t-1}(v)$. Then easy calculation leads to

$$A\big(\vec{\sigma}(A\boldsymbol{x}(t-1) + Bv) - \vec{\sigma}(A\boldsymbol{x}(t-1) + Bu(t-1))\big) = Bu$$

and hence $B \in \operatorname{image} A$ which contradicts $\operatorname{rank}[A, B] = 2$. Here we used the assumption that $\operatorname{rank} A < 2$. It remains to show that γ is not tangent

to c_{t-1}. But this follows by the observation that small pertubation of the input $u(t-1)$ would yield to a curve $\tilde{\gamma}$ which crosses c_{t-1} at least two times which is not possible. Of course it is clear from Lemma 3.1 that there is at most one curve γ which is tangent to c_{t-1} and claim 2 is proved.
To prove claim 1 observe that

$$\lim_{v\to\pm\infty} c_{t-1}(v) = \lim_{v\to\pm\infty} c_0(v)$$

and therefore any curve crossing c_{t-1} also crosses c_0, which proves the proposition.

□

References

1. F. Albertini, V. Maillot, and E.D. Sontag, *Uniqueness of weights for neural networks*, pp. 115–125, Chapman and Hall, London, 1993.
2. F. Albertini and P. Dai Pra, *Forward accessibility for recurrent neural networks*, IEEE Trans. Autom. Control **40** (1995), 1962–1968.
3. ———, *Recurrent neural networks: Identification and other system theoretic properties*, Neural Network Systems Techniques and Applications **3** (1997), 1–41.
4. M. Deistler and M. Dörfler, *A structure theory for identification of recurrenet neural networks – part 1*, Proceedings of the 4th IFAC "Nonlinear Control Systems Design Symposium" (1998), 459–464.
5. Y. Qiao and E.D. Sontag, *Further results on controllability of recurrent networks*, Systems and Control Letters **36** (1999), 121–129.
6. E.D. Sontag, *Mathematial control theory*, TAM 6 Springer Verlag, Berlin Heidelberg New York, 1990.
7. E.D. Sontag and H.J. Sussmann, *Complete controllability of continuous–time recurrent networks*, Systems and Control Letters **30** (1997), 177–183.
8. Thomas Steinberger and Lucas Zinner, *Complete controllability of discrete–time recurrent neural networks*, (1999), 1–5.

E-mail address: Thomas.Steinberger@@tuwien.ac.at and Lucas.Zinner@@tuwien.ac.at

Institute of Econometrics, Operations Research and System Theory, Technical University Vienna, Austria
E-mail address: Thomas.Steinberger@@tuwien.ac.at and Lucas.Zinner@@tuwien.ac.at