Recursive Construction of Periodoc Steady State for Neural Networks

Martin Matamala

Laboratoire de l'Informatique du Parallélisme

Ecole Normale Supérieure de Lyon
Unité de recherche associée au CNRS n°1398

Martín Matamala

July 1993

Research Report N° 93-23

Ecole Normale Supérieure de Lyon
46, Allée d'Italie, 69364 Lyon Cedex 07, France,
Téléphone : + 33 72 72 80 00; Télécopieur : + 33 72 72 80 80;
Adresses électroniques :
lip@frensl61.bitnet; lip@lip.ens−lyon.fr (uucp).

Recursive Construction of Periodic Steady State for Neural Networks

Martín Matamala

July 1993

Abstract

We present a strategy in order to build neural networks with long steady state periodic behavior. This strategy allows us to obtain 2^n non equivalent neural networks of size n, when the equivalence relation is the dynamical systems one. As a particular case, we build a neural network with n neurons admitting a cycle of period 2^n.

Keywords: neural networks, dynamical systems

Résumé

Nous presentons une stratégie pour construir des reseaux neuronaux qui ont cycles de grand taille. Cette stratégie nous permets d'obtenir 2^n reseaux neuronaux de taille n, qui sont no équivalents pour la relation de équivalence habituelle dans les systèmes dynamiques. Comme un cas particulier, nous construisons un reseaux neuronal de taille n, qui a un seul cycle de longueur 2^n.

Mots-clés: reseaux neuronaux, systèmes dynamiques

Recursive Construction of Periodic Steady State for Neural Networks*

Martín Matamala†

Laboratoire de l'Informatique du Parallélisme

École Normale Supérieure de Lyon

46, Allée d'Italie, 69364 Lyon Cedex 07, France.

July 21, 1993

Abstract

We present a strategy in order to build neural networks with long steady state periodic behavior. This strategy allows us to obtain 2^n non equivalent neural networks of size n, when the equivalence relation is the dynamical systems one. As a particular case, we build a neural network with n neurons admitting a cycle of period 2^n.

1 Introduction

A neural network of size n, is a discrete dynamical system acting on $\{-1,1\}^n$, whose transition function, F_A, is given in term of an nxn real matrix $A=(a_{ij})$ as follows:

$$F_A(x)=\overline{sgn}(Ax); \quad (Ax)_i=\sum_{j=1}^{n} a_{ij}x_j \quad i=1,...,n$$

$$\overline{sgn}: R^n \to \{-1,1\}^n, \quad \overline{sgn}(y)_i = sgn(y_i) \ i=1,...,n \tag{1}$$

* Partially supported by FONDECYT 0047/90

†Adress: From December 93: Departamento de Ingeniería Matemática, U. de Chile, Casilla 170-correo 3 Santiago, Chile

$$sgn(u) = \begin{cases} 1 & u \geq 0 \\ -1 & u < 0 \end{cases}$$

We are interested in the reverberation neural networks, i.e. neural networks where each state of the system, after a finite number of steps comes back to himself (hypercube permutations). We study the question: *how many reverberation neural networks have really different dynamical behavior ?*. A related question is asked in [2] where an equivalence relation is defined whose equivalence classes are characterized and the number of elements in any class is proved to be $2^n n!$ where n is the size of the neural network, but nothing is said concerning the number of different classes.

Our approach is slightly different and it consists in considering the neural networks , in particular reverberation neural networks, as dynamical systems. In order to give a partial answer to our question, we define also an equivalence relation and we prove, by building recursively 2^n non equivalent neural networks, that there exists at least 2^n different classes when we consider this relation in the set of the reverberation neural networks of size n.

This work is divided in two parts: the recursive construction of neural networks and the enumeration of different dynamical behaviors.

In the former, we give a process which permits us to build recursively neural networks satisfying two properties: *strictness* and *variability* which are weak enough so that one can find a large number of neural networks satisfying them. This process is supported by lemma 1 and 2; lemma 1 gives a way to build from a threshold function $f : \{-1,1\}^n \to \{-1,1\}$ another threshold function $g : \{-1,1\}^{n+1} \to \{-1,1\}$ such that over a vector (x,u) belongs to $\{-1,1\}^n \text{x} \{-1,1\} \backslash \{y_1,y_2,y_3,y_4\}$, $g(x,u) = f(x)$ and $g(y_i)$ for $i = 1,2,3,4$ is fixed by the construction. So, one can describe easily the dynamical evolution of g in term of those of f. Lemma 2, gives a way in order to build threshold functions which have an *a priori* desired behavior.

From these two lemmas we give in theorem 1 a recursive way for the construction of matrices , i.e., given a matrix A of size n satisfying hypotheses (a′) and (b′), defined in the next section, we build two matrices B and C of size $n+1$, satisfying also (a′) and (b′) . This process will permit us to find a large number of neural network which are defined by the matrices given in theorem 1.

In the second part we define an equivalence relation on P_n, the set of bijective functions from $\{-1,1\}^n$ into $\{-1,1\}^n$ and we build a function η associating to each element in P_n a vector of size 2^n. We prove that this function characterizes the equivalence relation, i.e., two functions F and G are equivalent iff $\eta(F) = \eta(G)$. Hence, we prove that the extensions B and C given in theorem 1 define non equivalent neural networks by proving that $\eta(F_B)$ and $\eta(F_C)$ are differents, where F_B (resp. F_C) is the transition function associated a B (resp. C). Latter, we prove that given two non equivalent neural networks A and A' their extensions are also non equivalent. This fact implies that increasing the size of the neural networks by one unit, one can double the non equivalent neural network number. That explains why we find 2^n non equivalent neural networks.

As a corollary, we build a neural network A of size n which has only one cycle of period 2^n.

2 Recursive construction of neural network

The following properties are important in our construction and represent the possibility of modification for a vector.

Definition 1 Consider $a \in R^n$. We say that

(a) a is *strict* if $\forall x \in \{-1,1\}^n \quad a \cdot x = \sum_{j=1}^{n} a_j x_j \neq 0$

(b) a is *variable* if $\exists I \in \{-1,1\}^n,\ a \cdot I < 0$ such that

$$\forall x \in \{-1,1\}^n \ x \neq I \quad a \cdot x < 0 \Rightarrow a \cdot x < a \cdot I$$

Observe that for a vector a satisfying (a) and (b) we have the scheme given in figure 1 which we adopt in order to give a more clear vision of the results.

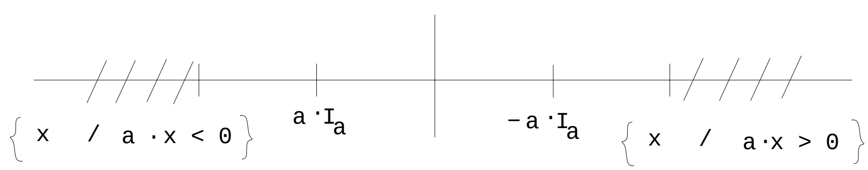

Figure 1: Scheme of the values of $a \cdot x$ where $x \in \{-1,1\}^n$ and a is a *strict variable* vector.

For a vector a satisfying (a) and (b) one can have only x such that $a \cdot x < 0$ or $a \cdot x > 0$ and then between $a \cdot I_a$ and $-a \cdot I_a$, in figure 1, there no exist any value $a \cdot x$.

Definition 2 Previous definitions apply to a real nxn matrix A by imposing that each row of A satisfies them. More precisely, given a matrix A we say :

(a′) A is *strict* if each row of A, a^i, for $i = 1, ..., n$, is strict.

(b′) A is *variable* if there exists a vector I_A such that a^i satisfies (b) with $I = I_A$, for every $i = 1, ..., n$.

When there exists a vector I_A (resp. I_a) satisfying (b′)(resp. b) we say that A(resp. a) is I_A(resp. I_a)-*variable*.

In the sequel we will work with vectors and matrices verifying properties (a) and (b). So, we define

$$M_n^*(R) = \{A : A \text{ is a } \textit{strict variable } n\text{x}n \text{ real matrix}\}$$

$$R_*^n = \{a \in R^n, \text{ is a } \textit{strict variable} \text{ vector}\}$$

Definition 3 The transition function F_a associated to vector $a \in R^n$ is given by:

$$F_a : \{-1, 1\}^n \to \{-1, 1\}$$

$$x \to F_a(x) = sgn(a \cdot x)$$

Observe that F_A given in equation 1 can be written as follows:

$$F_A(x)_i = F_{a^i}(x) = sgn(a^i \cdot x) \quad i = 1, ..., n$$

where a^i is the ith row of A.

The following lemmas give the vector basic extensions. In this lemmas several technical details are concentrated and in the sequel only its conclusions will be used.

In lemma 1 we build from a vector $a \in R_*^n$ another vector $b \in R_*^{n+1}$ such that the function F_b is an extension of the function F_a from $R^n \backslash \{I_a, -I_a\}$ to $R^{n+1} \backslash \{(\mu I_a, u)/\mu, u \in \{-1, 1\}\}$ and such that F_b over $\{(\mu I_a, u)/\mu, u \in \{-1, 1\}\}$ takes values depending only in the $n+1$th coordinate. In order to get a better understanding of lemma 1 we show the meaning of the concepts used on, by giving an example.

Consider the vector $a \in R^2$ given by

$$a^t = (1, -\frac{1}{2}) \tag{2}$$

Compute the values $a \cdot x$ for $x \in \{-1,1\}^2$. Since $a \cdot x = -a \cdot (-x)$ we get:

$$\begin{array}{rl} a \cdot (1,1) = 1 - \frac{1}{2} = \frac{1}{2} & a \cdot (-1,-1) = \dfrac{-1}{2} \\ a \cdot (-1,1) = -1 - \frac{1}{2} = -\frac{3}{2} & a \cdot (1,-1) = \dfrac{3}{2} \end{array}$$

Clearly a is *strict*. Let $I_a^t = (-1,-1)$. Then since

$$-\frac{3}{2} < -\frac{1}{2} < 0 < \frac{1}{2} < \frac{3}{2}$$

a is *I_a-variable.*

Let $D_n^-(a)$, $D_n^+(a)$ be given by

$$D_n^-(a) = \{x \in \{-1,1\}^n / a \cdot x < 0, \quad x \neq I_a\} \tag{3}$$

$$D_n^+(a) = \{x \in \{-1,1\}^n / a \cdot x > 0, \quad x \neq -I_a\} \tag{4}$$

In this case: $n = 2$, $I_a^t = (-1,-1)$ and $a^t = (1, -\frac{1}{2})$. Then $D_2^-(a) = \{(-1,1)\}$ $\quad D_2^+(a) = \{(1,-1)\}$. Let h_a be the maximum value in $D_n^-(a)$ given by:

$$h_a = \max\{a \cdot x : \ x \in D_n^-(a)\} \tag{5}$$

Then $h_a = -\frac{3}{2}$. Let $\delta > 0$ be such that $2a \cdot I_a = -1 < -\delta < -\frac{1}{2} = a \cdot I_a$ and $\frac{h_a + a \cdot I_a}{2} = -1 < -\delta < -\frac{1}{2} = a \cdot I_a$. Taking $\delta = \frac{3}{4}$ and $v = (I_a)_2 = -1$ we define $b \in R^3$ by

$$b_1 = a_1 = 1 \qquad b_2 = a_2 + v\frac{\delta}{2} = -\frac{1}{2} - \frac{3}{8} = -\frac{7}{8} b_{n+1} = -\frac{\delta}{2} = -\frac{3}{8}$$

Then b is as follows:

$$b^t = (1, -\frac{7}{8}, -\frac{3}{8}) \tag{6}$$

and $b \cdot (x,u)$ for $(x,u) \in \{-1,1\}^2 \mathrm{x} \{-1,1\}$ is given by (see figure 2)

$b\cdot(1,1,1) = -b\cdot(-1,-1,-1) = -\frac{1}{4}$ $b\cdot(-1,1,1) = -b\cdot(1,-1,-1) = -\frac{9}{4}$
$b\cdot(1,-1,1) = -b\cdot(-1,1,-1) = \frac{3}{2}$ $b\cdot(1,1,-1) = -b\cdot(-1,-1,1) = \frac{1}{2}$

For a vector $I \in R^n$ and an element $v \in R$ we denote by $(I,v)^t$ the extension of I from R^n into R^{n+1} whose $n+1$th coordinate is v.

So, b is *strict* and taking $I_b^t = (-I_a, 1) = (1,1,1)$ one obtains that $b\cdot I_b = -\frac{1}{4}$ and $h_b = -\frac{1}{2}$ and then b is I_b-*variable*. Moreover, F_b and F_a are related by:

$$F_b(-1,1,1) = F_b(-1,1,-1) = F_a(-1,1) = -1$$

$$F_b(1,-1,1) = F_b(1,-1,-1) = F_a(1,-1) = 1$$

and $F_b(I_a,1) = -1 = F_b(-I_a,1)$, $F_b(I_a,-1) = 1 = F_b(-I_a,-1)$. So,

$$F_b(x,u) = F_a(x) \quad x \neq I_a,\ x \neq -I_a \tag{7}$$

and

$$F_b(\mu I_a, u) = -u \quad \mu, u \in \{-1,1\} \tag{8}$$

Taking

$$\bar{a}^t = (-\frac{1}{2}, 1) \tag{9}$$

we deduce that $I_{\bar{a}}^t = (-1,-1)$, $h_{\bar{a}} = -\frac{3}{2}$ and $\bar{a}\cdot I_{\bar{a}} = -\frac{1}{2}$. For $\bar{a}$ we define vector $\bar{b}$ by:

$$\bar{b}^t = (-\frac{1}{2}, 1+(-1)\frac{3}{8}, -\frac{3}{8}) = (-\frac{1}{2}, \frac{5}{8}, -\frac{3}{8}) \tag{10}$$

It is easy to see that $\bar{b}$ is *strict*, $I_{\bar{b}}^t = (-I_a, v) = (1,1,1)$-*variable* and that $F_{\bar{b}}$ satisfies equations 7 and 8. The generalization of this result is given in lemma 1.

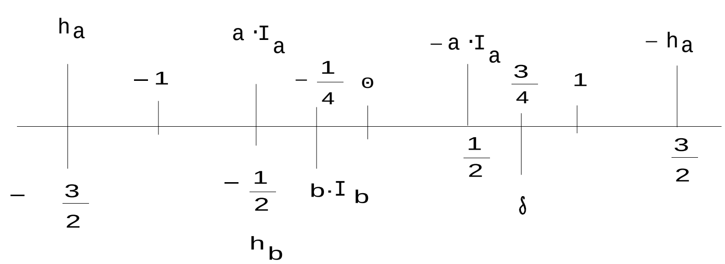

Figure 2: Scheme of the values of $(1, \frac{-1}{2})\cdot x$ where $x \in \{-1,1\}^2$.

Lemma 1. Let $a \in R^n_*$. Then, there exists $b \in R^{n+1}_*$ satisfying:

(a) $I^t_b = (-I_a, 1)$,

(b) $\forall (x, u) \in \{-1, 1\}^n \mathrm{x} \{-1, 1\} \quad x \neq -I_a,\ x \neq I_a \ \ F_b(x, u) = F_a(x)$

(c) $\forall \mu, u \in \{-1, 1\}\ F_b(\mu I_a, u) = -u$.

Proof

Let $D^-_n(a)$, $D^+_n(a)$ be given as in equations 3 and 4.

Since a is a *strict* vector we have the following equivalence

$$x \in \{-1, 1\}^n \text{ iff } x \in \{-I_a, I_a\} \ \vee x \in D^-_n(a) \ \vee x \in D^+_n(a) \tag{11}$$

Let h_a be the maximum value in $D^-_n(a)$ given in equation 5. Since a is I_a-*variable* we have that $h_a < a \cdot I_a < 0$ and then (see figure 3):

$$2a \cdot I_a < a \cdot I_a \text{ and } \frac{h_a + a \cdot I_a}{2} < a \cdot I_a \tag{12}$$

so, there exists $\delta > 0$ such that

$$2a \cdot I_a < -\delta < a \cdot I_a \text{ and } \frac{h_a + a \cdot I_a}{2} < -\delta < \tag{13}$$

which is equivalent to:

$$-(a \cdot I_a + \delta) < 0\ ,\ h_a + \delta < -(a \cdot I_a + \delta) \text{ and } a \cdot I_a < -(a \cdot I_a + \delta) \tag{14}$$

Observe that in figure 3 we suppose that $\frac{h_a + a \cdot I_a}{2} > 2a \cdot I_a > h_a$ which is not the general case. From definition of h_a and equation 14 we have for $x \in D^-_n(a)$ that:

$$a \cdot x + \delta \leq h_a + \delta < -(a \cdot I_a + \delta) < 0 \tag{15}$$

hence, $\delta < |a \cdot x|$. Since $x \in D^-_n(a)$ iff $-x \in D^+_n(a)$ we obtain

$$\forall x \in D^-_n(a) \cup D^+_n(a) \ \ |a \cdot x| > \delta \tag{16}$$

Define $b \in R^{n+1}$ by

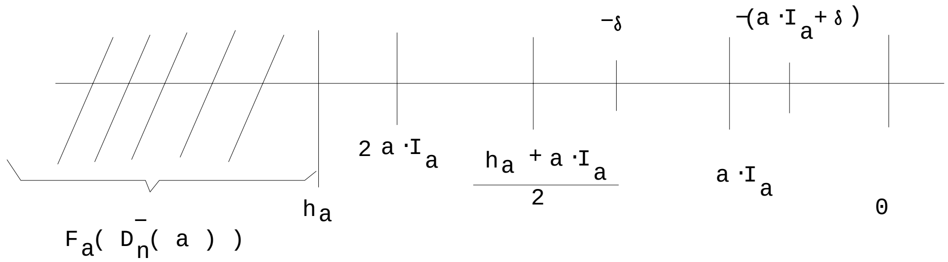

Figure 3: Scheme of the values of differents parameters defined in lemma 1.

$b_i = a_i \quad i = 1, ..., n-1 \qquad b_n = a_n + v\frac{\delta}{2} \qquad b_{n+1} = -\frac{\delta}{2}$

where $(I_a)_n = v$. It is clear that for $(x, u) \in \{-1, 1\}^n \mathrm{x} \{-1, 1\}$ we have

$$b \cdot (x, u) = a \cdot x + (v x_n - u)\frac{\delta}{2} \tag{17}$$

and then

$$|b \cdot (x, u) - a \cdot x| \leq \delta \tag{18}$$

We shall prove that b is *variable* . For that we want to find (x, u) satisfying $b \cdot (x, u) < 0$. Let $x \in D_n^-(a) \cup D_n^+(a)$ then from equations 16 and 18 we have:

$$a \cdot x > 0 \;\Rightarrow b \cdot (x, u) \geq a \cdot x - \delta > 0 \tag{19}$$

and

$$a \cdot x < 0 \;\Rightarrow b \cdot (x, u) \leq a \cdot x + \delta < 0 \tag{20}$$

Let $x = \mu I_a$ with $\mu = -1, 1$. Since $(I_a)_n = v$, from equation 17 we have:

$$b \cdot (\mu I_a, 1) = a \cdot \mu I_a + (vv\mu - 1)\frac{\delta}{2} = a \cdot \mu I_a + (\mu - 1)\frac{\delta}{2} \tag{21}$$

so, from equations 14 and 21 we obtain:

$$\mu = 1 \Rightarrow b \cdot (I_a, 1) = a \cdot I_a < -(a \cdot I_a + \delta) < 0 \tag{22}$$

$$\mu = -1 \Rightarrow b \cdot (-I_a, 1) = -a \cdot I_a + (-1 - 1)\tfrac{\delta}{2} = -(a \cdot I_a + \delta) < 0 \tag{23}$$

i.e., $b \cdot (\mu I_a, 1) < 0$ for $\mu = -1, 1$. Let $I_b = (-I_a, 1)$. Then from equation 23 we get $b \cdot I_b < 0$ and applying equations 19, 20, 22 and 23 we obtain the following equivalence:

$$b \cdot (x,u) < 0 \text{ and } (x,u) \neq I_b \text{ iff } a \cdot x < 0 \wedge x \neq I_a \vee (x,u) = (I_a, 1) \quad (24)$$

Let $(x,u) \in \{-1,1\}^{n+1}$ such that $b \cdot (x,u) < 0 \wedge (x,u) \neq I_b$. From equation 24 there are only two possibilities for (x,u). For the first one i.e, $a \cdot x < 0$ and $x \neq I_a$ we know, from equations 15, 20 and 23 that

$$b \cdot (x,u) \leq a \cdot x + \delta < -(a \cdot I_a + \delta) = b \cdot I_b$$

For the second one, from equation 22 we get that

$$b \cdot (I_a, 1) = a \cdot I_a < b \cdot I_b$$

which proves that b is I_b-*variable*. Observe that the inequalities in equations 19, 20, 22 and 23 are strict, so,

$$\forall (x,u) \in \{-1,1\}^{n+1} \quad b \cdot (x,u) < 0 \vee b \cdot (x,u) > 0$$

which says that b is *strict*.

Finally, from equations 19, 20, 22 and 23 we get:

$$\forall (x,u) \in \{-1,1\}^{n+1} \; x \in D_n^-(a) \cup D_n^+(a)$$

$$F_b(x,u) = sgn(b \cdot (x,u)) = sgn(a \cdot x) = F_a(x)$$

and

$$\forall u, \mu \in \{-1,1\} \quad F_b(\mu I_a, u) = sgn(b \cdot (\mu I_a, u)) = sgn(-u) = -u$$

■

In next lemma we build two vectors c and d in R_*^{n+1}. Vector d is such that the function F_d is the projection over the $n+1$th coordinate. Vector c defines the function F_c being the projection of the $n+1$th coordinate from $R \backslash \{(\mu I_a, u)/\mu u \in \{-1,1\}\}$ into $\{-1,1\}$ and it considers only the sign of the nth coordinate of $(\mu I_a, u)$ for $\mu, \; u \in \{-1,1\}$.

In order to show how proceeds the proof of lemma 2 we give an example:

Let $I_a = (-1, -1)$. We define

$$c = (-1, -1, 2 - \frac{1}{2}) = (-1, -1, \frac{3}{2}) \tag{25}$$

and

$$d = (-1, -1, 2 - \frac{-1}{2}) = (-1, -1, \frac{5}{2}) \tag{26}$$

Then $c \cdot (x, u)$ for $(x, u) \in \{-1, 1\}^2 \text{x} \{-1, 1\}$ is given by:

$$c \cdot (-1, -1, -1) = -c \cdot (1, 1, 1) = \frac{1}{2} \quad c \cdot (-1, -1, 1) = -c \cdot (1, 1, -1) = \frac{7}{2}$$

$$c \cdot (-1, 1, 1) = -c \cdot (1, -1 - 1) = \frac{3}{2} \quad c \cdot (-1, 1, -1) = -c \cdot (1, -1, 1) = -\frac{3}{2}$$

so, c is *strict*, $I_c = (-I_a, 1) = (1, 1, 1)$-*variable* and satisfies $F_c(x, u) = u \ x \neq I_a, \ x \neq -I_a$, $F_c(-I_a, -1) = -1$ and $F_c(-I_a, 1) = 1$ i.e., $F_c(\mu I_a, u) = \mu$.

By the other hand $d \cdot (x, u)$ is given by:

$$d \cdot (-1, -1, -1) = -d \cdot (1, 1, 1) = -\frac{1}{2} \quad d \cdot (-1, -1, 1) = -d \cdot (1, 1, -1) = \frac{9}{2}$$

$$d \cdot (-1, 1, 1) = -d \cdot (1, -1 - 1) = \frac{5}{2} \quad d \cdot (-1, 1, -1) = -d \cdot (1, -1, 1) = -\frac{5}{2}$$

so, d is *strict*, $(I_a, -1)$-*variable* and $F_d(x, u) = u$

Lemma 2. For $I_a \in R^n, (I_a)_n = v$ there exist c and $d \in R_*^{n+1}$ such that:

(a) $I_c = (-I_a, 1) \ I_d = (I_a, -1)$

(b.1) $\forall (x, u) \in \{-1, 1\}^n \text{x} \{-1, 1\} \quad x \neq I_a \ x \neq -I_a \ \ F_c(x, u) = u$

(b.2) $\forall \mu, u \in \{-1, 1\} \ \ F_c(\mu I_a, u) = \mu$.

(c) $\forall (x, u) \in \{-1, 1\}^n \text{x} \{-1, 1\} \ F_d(x, u) = u$

Proof The construction of c and d is very similar. So, we give this construction in only a vector $e(r)$ which will be appropriately evaluated in order to obtain c and d.

Let $e(r)^t = \left(I_a, (n - \frac{r}{2})\right), \ |r| = 1$ belong to R^{n+1}.

For $(x,u) \in \{-1,1\}^n$x$\{-1,1\}$ we have

$$e(r)\cdot(x,u) = x\cdot I_a + u(n-\frac{r}{2}) \tag{27}$$

It is easy to see that $x \neq \mu I_a$, $\mu=-1,1$ is equivalent to $-n+2 \leq x\cdot I_a \leq n-2$, which applied to equation 27 implies:

$$u=1 \quad e(r)\cdot(x,1) \geq -n+2+n-\frac{r}{2} = 2-\frac{r}{2} > 1 \tag{28}$$

$$u=-1 \quad e(r)\cdot(x,-1) \leq -n+\frac{r}{2}+n-2 = -2+\frac{r}{2} < -1 \tag{29}$$

i.e.,

$$sgn(e(r)\cdot(x,u)) = u \text{ when } x \neq \mu I_a,\ \mu=-1,1$$

Let $x=\mu I_a$, $\mu=1,-1$. Then

$$e(r)\cdot(\mu I_a,u) = \mu I_a\cdot I_a + u(n-\frac{r}{2}) = (\mu+u)n - \frac{ur}{2}$$

So, $|e(r)\cdot(\mu I_a,u)| = |(\mu+u)n-\frac{ur}{2}| \geq \Big|\ |\mu+u|n - |\frac{ur}{2}|\ \Big| = \Big|\ |\mu+u|n-\frac{1}{2}\ \Big| > 0$. Hence, from equations 28 and 29 $e(r)$ is *strict.*

We prove that $e(r)$ is $I_{e(r)}$-*variable* with $I^t_{e(r)} = r(-I_a,1)$. Compute the value $e(r)\cdot I_{e(r)}$:

$$e(r)\cdot I_{e(r)} = \left((-r+r)n+-\frac{rr}{2}\right) = -\frac{1}{2}$$

moreover,

$$e(r)\cdot(-I_a,-1) = (-2n+\frac{r}{2}) = -2n+\frac{r}{2} < -1$$

Since $r(-I_a,1) \neq (-I_a,-1)$ we obtain:

$(x,u) \neq I_{e(r)}$ and $e(r)\cdot(x,u) < 0$ iff $x \neq I_a$, $x \neq -I_a$ and $u=-1$ or $(x,u)=(-I_a,-1)$ and then for $e(r)\cdot(x,u)<0$ we get

$$(x,u) \neq I_{e(r)} \Rightarrow e(r)\cdot(x,u) < -1 < e(r)\cdot I_{e(r)} = -\frac{1}{2}$$

So, $e(r)$ is $I_{e(r)}$-*variable.*

Taking $c=e(1)$ and $d=e(-1)$ it is easy to see that (a), (b) and (c) are satisfied.

■

The extension for a matrix A is given in the following theorem. As an example of the construction consider the real matrix A given by:

$A = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} a \\ \bar{a} \end{pmatrix}$ where a and $\bar{a}$ are given in equations 2 and 9. Then from the analysis for a and $\bar{a}$, A is *strict* and $I_A = (-1,-1)$-*variable*. Consider B given by $B = \begin{pmatrix} b \\ \bar{b} \\ c \end{pmatrix} = \begin{pmatrix} 1 & -\frac{7}{8} & \frac{-3}{8} \\ -\frac{1}{2} & \frac{5}{8} & \frac{-3}{8} \\ -1 & -1 & \frac{3}{2} \end{pmatrix}$ where B, $\bar{b}$ and c are constructed by equations 6, 10 and 25. Then B is *strict* and $(1,1,1)$-*variable*. Moreover, for $x \neq I_A,\ x \neq -I_A$

$$\begin{aligned} F_B(x,u) &= (F_b(x,u), F_{\bar{b}}(x,u), F_c(x,u))^t \\ &= (F_a(x), F_{\bar{a}}(x), u)^t = (F_A(x), u)^t \end{aligned}$$

and

$$\begin{aligned} F_B(\mu I_A,u) &= (F_b(\mu I_A,u), F_{\bar{b}}(\mu I_A,u), F_c(\mu I_A,u))^t \\ &= (-u,-u,\mu)^t = (-ue_2,\mu)^t \end{aligned}$$

where $e_2^t = (1,1)$. Now, let $C = \begin{pmatrix} a & 0 \\ \bar{a} & 0 \\ d & \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & 0 \\ -1 & -1 & \frac{5}{2} \end{pmatrix}$ where d is given by equation 26. Then C is *strict* and $(I_A,-1)$-*variable*. Moreover,

$$F_C(x,u) = (F_a(x), F_{\bar{a}}(x), F_d(x,u))^t = (F_A(x), u)^t$$

The last construction is generalized for any matrix in $M_n^*(R)$ in the following theorem.

Theorem 1. For $A \in M_n^*(R)$, there exist B and C in $M_{n+1}^*(R)$ such that

(a) $I_B = (-I_A, 1), \quad I_C = (I_A, -1)$

(b) $\forall (x,u) \in \{-1,1\}^n \mathrm{x} \{-1,1\} \quad F_C(x,u) = (F_A(x),u)^t$

(c) $\forall (x,u) \in \{-1,1\}^n \mathrm{x} \{-1,1\} \quad x \neq -I_A,\ x \neq I_A \quad F_B(x,u) = (F_A(x),u)^t$

$$\forall \mu, u \in \{-1,1\} \quad F_B(\mu I_A, u) = (-ue_n, \mu)^t \tag{30}$$

where $e_n^t = (1, ..., 1) \in R^n$.

Proof

Construction of matrix B :

Since $A \in M_n^*(R)$ we know that each row, a^i, $i = 1, ..., n$ belongs to R_*^n. Applying Lemma 1 to each row we find vectors $b^i \in R_*^{n+1}$ which are $(-I_A, 1)$-*variables*, satisfying:

$\forall (x, u) \in \{-1,1\}^n \mathrm{x} \{-1,1\} \quad x \neq -I_A,\ x \neq I_A \quad F_{b^i}(x, u) = F_{a^i}(x)$ and $F_{b^i}(\mu I_A, u) = -u$

By applying Lemma 2, for $I = I_A$, we obtain $c \in R^{n+1}$, $(-I_A, 1)$-*variable*, such that

$\forall (x, u) \in \{-1,1\}^n \mathrm{x} \{-1,1\} \quad x \neq -I_A,\ x \neq I_A \ F_c(x, u) = u$ and $F_c(\mu I_A, u) = \mu$

Define $B^t = (b^1, b^2, ..., b^n, c)$. Since each b^i, for $i = 1, ..., n$ belongs to R_*^{n+1} with $I_{b^i} = (-I_A, 1)$ one knows that $B \in M_{n+1}^*(R)$ and from lemma 1 and conclusion (b) of lemma 2, B verifies properties (a) and (c) of the theorem .

Construction of matrix C :

Let d be the vector given by lemma 2 which is $(I_a, -1)$-variable. Let C be defined by:

$$C_{ij} = a_{ij} \quad 1 \leq i, j \leq n \quad C_{j,n+1} = 0 \quad 1 \leq j \leq n \quad C_{n+1,j} = d_j \quad 1 \leq j \leq n+1$$

Since $d \in R_*^{n+1}$ with $I_d = (I_A, -1)$, $C \in M_{n+1}^*$. Moreover, $C \begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} Ax \\ d \cdot u \end{pmatrix}$ and then F_C is given by $F_C(x, u) = (F_A(x), F_d(u))^t = (F_A(x), u)$. So, C satisfies (a) and (b) in theorem. ■

3 Non equivalent neural networks

Consider the set P_n of the bijective functions on $\{-1,1\}^n$. The following property is shown in [2] for $F \in P_n$.

$$\forall x \in \{-1,1\}^n \ \exists s \in I\!N \quad F^s(x) = x \tag{31}$$

We define the cycle of x by F, $O_F(x)$, for $F \in P_n$ by

$$O_F(x) = < x, F(x), ..., F^{T_x^F - 1}(x) >$$

where T_x^F is the first integer such that $F^{T_x^F}(x) = x$. T_x^F is called the period of the cycle $O_F(x)$. We say that $y \in O_F(x)$ iff there exists $s \in I\!N$ such that $F^s(x) = y$. Taking $A = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$ we have

$$O_{F_A}(1,1) =< (1,1) > \qquad O_{F_A}(-1,-1) = - < (-1,-1) >$$

and

$$O_{F_A}(1,-1) =< (1,-1), (1,-1) >$$

Then

$$T_{(1,1)}^{F_A} = 1 = T_{(-1,-1)}^{F_A}, \;\; T_{(-1,1)}^{F_A} = T_{(1,-1)}^{F_A} = 2$$

In order to show the power of the construction given in section 2 it is necessary to specify when two neural networks have different dynamics. For that we define the following equivalence relation:

Given F and G in P_n we say that F is equivalent to G iff there exists a function Φ on P_n such that

$$\forall x \in \{-1,1\}^n \quad F(\Phi(x)) = \Phi(G(x)) \tag{32}$$

This definition does not permit easily to prove that our construction builds non equivalent neural network. For that, given a function $F \in P_n$, we define the characteristic of F by a vector $\eta(F)$ in $I\!N^{2^n}$, such that its ith component gives the cycle numbers of period i of F and we prove the following lemma.

Lemma 4. Two functions F and G in P_n are equivalent iff $\eta(F) = \eta(G)$.

Proof ($\Rightarrow$) We prove the following equivalence for Φ satisfying equation 32:

$$C^F =< x, F(x), ..., F^{L-1}(x) > \text{ is a cycle for} F \text{ iff} \tag{33}$$

$C^G =< \Phi(x), \Phi(F(x)), ..., \Phi(F^{L-1}(x)) >$ is a cycle for G.

Indeed, since F and G are equivalent we have that

$$\Phi(F^i(x)) = G^i(\Phi(x)) \text{ for } i = 0, ..., L-1$$

and then equation 32 is true. Hence for each cycle of size L of F we have a cycle of size L for G and conversely for each cycle of size L of G we have a cycle of size L for G with which $\eta(F) = \eta(G)$.

($\Leftarrow$) Since F and G belong to P_n, a vector $x \in \{-1,1\}^n$ can belong to only one cycle.

Let C^i_j, β^i_j $j = 1, ..., n_i$ be the different cycles of size i for F and G respectively. We define the function Φ associating C^i_j to β^i_j as follows:

Let $C^i_j =< x, F(x), ..., F^{i-1}(x) >$ and $\beta^i_j =< y, G(y), ..., G^{i-1}(y) >$ then we define Φ by:

$$\Phi(G^k(y)) = F^k(x) \;\; 1 \leq k \leq i-1 \;\; \Phi(y) = x$$

Making this process for any j and any i we define completely Φ satisfying $\Phi F = G\Phi$. ■

Definition We say that a real matrix A is a reverberation neural network if F_A belong to P_n.

Proposition 1. Let $A \in M^*_n(R)$ a reverberation neural network. Then B and C given in theorem 1 are reverberation neural networks and the periods of their cycles are determined in terms of the periods of the cycles of F_A as follows:

$$\forall (x,u) \in \{-1,1\}^n \text{x} \{-1,1\} \qquad T^{F_C}_{(x,u)} = T^{F_A}_x \tag{34}$$

$$\forall (x,u) \in \{-1,1\}^n \text{x} \{-1,1\} \quad I_A, -I_A \notin O_{F_A}(x) \quad T^{F_B}_{(x,u)} = T^{F_A}_x \tag{35}$$

$$\text{if } \mu I_A \in O_{F_A}(x) \;\text{ for some } \mu = -1, 1 \text{ then } \; T^{F_B}_{(x,u)} = T^{F_B}_{(\mu I_A, u)} \tag{36}$$

$$\text{where } T^{F_B}_{(\mu I_A,u)} = \begin{cases} 2T^{F_A}_{I_A} & \text{if } e_n \in O_{F_A}(I_A) \\ 2T^{F_A}_{I_A} & \text{if } u = -\mu \text{ and } e_n \notin O_{F_A}(I_A) \\ T^{F_A}_{I_A} & \text{if } u = \mu \text{ and } e_n \notin O_{F_A}(I_A) \end{cases} \tag{37}$$

Proof Before giving the proof we analyze our example: From the definition of B and C in it is easy to see that:

$$O_{F_B}(1,-1,-1) = O_{F_B}(-1,1,-1) =< (1,-1,-1), (-1,1,-1) >$$

$$O_{F_B}(-1,1,1) = O_{F_B}(1,-1,1) =< (-1,1,1),(1,-1,1) >$$

$$O_{F_B}(-1,-1,1) = -O_{F_B}(1,1,-1) =< (-1,-1,1) >$$
$$O_{F_B}(1,1,1) = O_{F_B}(-1,-1,-1) =< (1,1,1),(-1,-1,-1) >$$

and

$$O_{F_C}(1,-1,-1) = O_{F_C}(-1,1,-1) =< (1,-1,-1),(-1,1,-1) >$$
$$O_{F_C}(-1,1,1) = O_{F_C}(1,-1,1) =< (-1,1,1),(1,-1,1) >$$

$$O_{F_C}(-1,-1,1) = -O_{F_C}(1,1,-1) =< (-1,-1,1) >$$
$$O_{F_C}(1,1,1) = -O_{F_C}(-1,-1,-1) =< (1,1,1) >$$

and then

$$\forall (x,u) \in \{-1,1\}^2 \mathrm{x} \{-1,1\} \quad T^{F_C}_{(x,u)} = T^{F_A}_x$$
$$\forall x \in \{-1,1\}^2 \ x \neq I_A \ x \neq -I_A \ \ T^{F_B}_{(x,u)} = T^{F_A}_x$$
$$T^{F_B}_{(I_A,1)} = T^{F_B}_{(-I_A,-1)} = 1 = T^{F_A}_{I_A}$$

and

$$T^{F_B}_{(-I_A,1)} = T^{F_B}_{(I_A,-1)} = 2 = 2T^{F_A}_{I_A}$$

Note that we are in the case $e \notin O_{F_A}(I_A)$, $(I_A, 1)$ and $(-I_A, -1)$ satisfy the condition $u = \mu$ and $(-I_a, 1)$ and $(-I_A, -1)$ satisfy the condition $u = -\mu$. This proves the proposition in the our example.

Now we give the general proof. Firstly we prove that B and C are reverberation neural networks. Suppose that $F_C(x, u) = F_C(x', u')$. Since $F_C(x, u) = (F_A(x), u)$ we have that $u = u'$ and $F_A(x) = F_A(x')$. But A is a reverberation neural network, so $(x, u) = (x', u')$ and C is a reverberation neural network. Now, suppose that $F_B(x, u) = F_B(x', u')$. Then if $x \neq \mu I_A$ we proceed as above. When $x = \mu I_A$ we have $F_B(\mu x, u) = (-ue_n, \mu)$. Since A is a reverberation neural network, $F_A(y) = -ue_n$ only for $y = \mu I_A$. Then $x' = \mu' I_A$ and from $(-ue_n, \mu) = (-u'e_n, \mu')$ we conclude that $(x, u) = (x', u')$ and B is a reverberation neural network.

Properties 34 and 35 follow from the fact that

$\forall k \in I\!N \quad F^k_C(x, u) = (F^k_A(x), u)$ and $F^k_B(x, u) = (F^k_A(x), u)$ when $F^k_A(x) \neq -I_A, I_A$.

When $\mu I_A \in O_{F_A}(x)$ we have that

$$O_{F_B}(x,u) =< (x,u),...,(z,t),(\mu I_A,u),(-ue,\mu),...,(y,w) >$$

and since $F_B \in P_{n+1}$, $O_{F_B}(\mu I_A, u)$ is given by:

$$O_{F_B}(\mu I_A,u) =< (\mu I_A,u),(-ue,\mu),...,(y,w)(x,u),...,(z,t) >$$

Observe the structure of $O_{F_B}(\mu I_A, u)$. Suppose that $e \in O_{F_A}(I_A)$ then since $F_A \in P_n$ the following sequence of transition is true:

$$I_A \to -e \to \cdots \to -I_A \to \underbrace{e \to \cdots \to I_A}_{T_{I_A}^{F_A}}$$

and then from Theorem 1

$$(\mu I_A,\mu) \to (-\mu e,\mu) \to \cdots \to (-\mu I_A,\mu) \to (-\mu e,-\mu) \to \cdots \to (-\mu I_A,-\mu)$$

$$(-\mu I_A,-\mu) \to (\mu e,-\mu) \to \cdots \to (\mu I_A,-\mu) \to (\mu e,\mu) \to \cdots \to (\mu I_A,\mu)$$

i.e., $T^{F_B}_{(\mu I_A,\mu)} = T^{F_B}_{(\mu I_A,-\mu)} = T^{F_B}_{(-\mu I_A,\mu)} = T^{F_B}_{(-\mu I_A,-\mu)} = 2T^{F_A}_{I_A}$

If $e \notin O_{F_A}(I_A)$ then

$$(\mu I_A,\mu) \to (-\mu e,\mu) \to \cdots \to (\mu I_A,\mu)$$

i.e., $T^{F_B}_{(\mu I_A,\mu)} = T^{F_A}_{(\mu I_A)}$, moreover,

$$(\mu I_A,-\mu) \to (\mu e,\mu) \to \cdots \to (-\mu I_A,\mu)$$

$$(-\mu I_A,\mu) \to (-\mu e,-\mu) \to \cdots \to (\mu I_A,-\mu)$$

i.e.,$T^{F_B}_{(\mu I_A,-\mu)} = 2T^{F_A}_{(\mu I_A)}$ and we have the conclusions. ∎

Observe that in our example we have $\eta(F_A) = (2,2,0,0)$, $\eta(F_B) = (2,3,0,0,0,0,0,0)$ and $\eta(F_C) = (4,2,0,0,0,0,0,0)$ which motives the following corollary which is a conclusion of lemma 4 and proposition 1.

Corollary 1 For matrices A, B and C in proposition 1 we have

(a) $\eta(F_C)_i = 2\eta(F_A)_i \;\; 1 \le i \le 2^n \qquad \eta(F_C)_i = 0 \;\; 2^n < i \le 2^{n+1}$

(b) $\eta(F_B)_i = 2\eta(F_A)_i \;\; 1 \leq i \leq 2^n \;\; \eta(F_B)_i = 0 2^n < i \leq 2^{n+1} \quad i \neq T^A_{I_A} \; i \neq 2T^A_{I_A}$

If $e \in O_{F_A}(I_A)$ then
(c) $\eta(F_B)_{T^{F_A}_{I_A}} = 2(\eta(F_A)_{T^{F_A}_{I_A}} - 1)$

$$\eta(F_B)_{2T^{F_A}_{I_A}} = 1 + \begin{cases} 2\eta(F_A)_{2T^{F_A}_{I_A}} & 2T^{F_A}_{I_A} \leq 2^n \\ 0 & 2T^{F_A}_{I_A} > 2^n \end{cases}$$

if $e \notin O_{F_A}(I_A)$ then
(d) $\eta(F_B)_{T^{F_A}_{I_A}} = 2(\eta(F_A)_{T^{F_A}_{I_A}} - 2) + 2$

$$\eta(F_B)_{2T^{F_A}_{I_A}} = 1 + \begin{cases} 2\eta(F_A)_{2T^{F_A}_{I_A}} & 2T^{F_A}_{I_A} \leq 2^n \\ 0 & 2T^{F_A}_{I_A} > 2^n \end{cases}$$

Since $\eta(F_B)_{2T^{F_A}_{I_A}}$ is odd, the neural network B and C are not equivalent.

Proof Observe that since $2(\eta(F_A)_{T^{F_A}_{I_A}} - 1) = 2(\eta(F_A)_{T^{F_A}_{I_A}} - 2) + 2$ we could join (c) and (d). For sake of clearness we prefer this form.

From Proposition 1 one knows that from each cycle $O_{F_A}(x)$ we can obtain two cycles $O_{F_C}(x, -1)$ and $O_{F_C}(x, 1)$ with the same period and that the cycle numbers of a given size of F_A is doubled in F_C. This same argument is true for F_B when the cycle O_{F_A} does not contain neither I_A nor $-I_A$. When I_A or $-I_A$ belongs to $O_{F_A}(x)$ we know that if $e \in O_{F_A}(I_A)$ then $O_{F_A}(I_A) = O_{F_A}(-I_A)$ and the cycle, $O_{F_A}(I_A)$ which is of size $T^{F_A}_{I_A}$, is transformed in the cycle $O_{F_B}(\mu I_A, \mu)$ of size $2T^{F_A}_{I_A}$. This is described by (c). If $e \notin O_{F_A}(I_A)$ then $O_{F_A}(I_A) \neq O_{F_A}(-I_A)$ and both are transformed in the cycle $O_{F_B}(I_A, -1)$ of size $2T^{F_A}_{I_A}$, cycle $O_{F_B}(I_A, 1)$ of size $T^{F_A}_{I_A}$ and cycle $O_{F_B}(-I_A, -1)$ of size $T^{F_A}_{F_A}$. The last observations is trivial from the definition of η. ■

Proposition 2. Let $\{A^i\}_{i=1}^L$ be a family of non equivalent reverberation neural networks in $M^*_n(R)$. Then $\{B^i, C^i\}_{i=1}^L$ is a family of non equivalent reverberation neural networks in $M^*_{n+1}(R)$, where B^i and C^i are built from A^i in theorem 1.

Proof

Suppose that there exist two equivalent neural networks in $\{B^i, C^i\}_{i=1}^L$. Then, it is sufficient to analyze the following cases:

(a) $\eta(F_{B^i}) = \eta(F_{B^j})$. Then we have that $\forall 1 \leq k \leq 2^n$, $k \neq T_{I_A}^{F_A}$ and $k \neq 2T_{I_A}^{F_A}$

$$\eta(F_{B^i})_k = \eta(F_{B^j})_k \Rightarrow \eta(F_{A^i})_k = \eta(F_{A^j})_k$$

and from (c) and (d) in Corollary 1 one obtains $\eta(F_{A^i}) = \eta(F_{A^j})$

(b) $\eta(F_{C^i}) = \eta(F_{C^j})$. Applying the same arguments as in (a) we conclude that $\eta(F_{A^i}) = \eta(F_{A^j})$ so, (a) and (b) are in contradiction with the non equivalence of A_i and A_j.
(c) $\eta(F_{B^i}) = \eta(F_{C^j})$. Then

$$\eta(F_{B^i})_{2T_{A^i}^{A^i}} = 1 + \begin{cases} \text{even} & 2T_{I_{A^i}}^{A^i} \leq 2^n \\ 0 & 2T_{I_{A^i}}^{A^i} > 2^n \end{cases}$$

and

$$\eta(F_{C^j})_{2T_{A^i}^{A^i}} = \begin{cases} \text{even} & 2T_{A^i}^{A^i} \leq 2^n \\ 0 & 2T_{A^i}^{A^i} > 2^n \end{cases}$$

but, this is a contradiction too. ■

Theorem 2 For any $n \in \mathbb{N}$ there exist 2^n non equivalent reverberation neural networks in $M_n^*(\boldsymbol{R})$

Proof We proceed by induction on n

For $n = 2$ the matrices $A^i \; : \; i = 1, 2, 3, 4$ given by

$$A^1 = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} \quad A^2 = \begin{pmatrix} -\frac{1}{2} & 1 \\ 1 & -\frac{1}{2} \end{pmatrix} \quad A^3 = \begin{pmatrix} -1 & \frac{1}{2} \\ \frac{1}{2} & -1 \end{pmatrix} \quad A^4 = \begin{pmatrix} \frac{1}{2} & 1 \\ -1 & \frac{-1}{2} \end{pmatrix}$$

are in M_2^*, and have the following characteristics

$$\eta(A^1) = (4, 0, 0, 0) \;\; \eta(A^2) = (2, 1, 0, 0) \;\; \eta(A^3) = (0, 2, 0, 0) \;\; \eta(A^4) = (0, 0, 0, 1)$$

and then are not equivalent. Accepting that there exists 2^n non equivalent neural networks for matrices of size n we can apply proposition 2 in order to obtains 2^{n+1} non equivalent neural networks for size $n + 1$. ■

By using corollary 1 we get the following result which is given in [1]:

Corollary 2 $\forall n \in \mathbb{N}$ there exists $A \in M_n^*(R)$ whose characteristic is given by:

$\eta(F_{A_k})_i = 0 \quad \text{for } i \neq 2^n \qquad \text{and } \eta(F_{A_k})_{2^n} = 1$

Proof Taking $n = 2$ we have that A^4 given by theorem 2 belong to M_2^* and its characteristic is $(0, 0, 0, 1)$. Accepting that there exist $A \in M_n^*(R)$ with $\eta(A) = (0, ..., 1)$ then by corollary 1 we obtain $B \in M_{n+1}^*(R)$ with $\eta(B) = (0, ..., 1)$ because $e_n \in O_{F_A}(I_A)$. ■

4 Conclusion

The results shown in this work permit us to obtain a wide variety of non equivalent dynamics when we consider the family of reverberation neural networks in $M_n^*(R)$. This kind of constructions can be applied for information storage where the information is codified in the cycles of the neural network.

It is desirable to extend our construction to any function in $M_n^*(R)$. In this case theorem 1 is true and we can build recursively neural networks in $M_n^*(R)$. Moreover, we can obtain an analogous result to proposition 1 which permits us to know the behavior of neural networks of size $n + 1$ in term of those of the neural networks of size n. But, the characterization given in lemma 4 for the equivalence of two functions in $M_n^*(R)$ is not longer true. For that, it is interesting to find an invariant in the general case.

References

[1] S. A. Gustafson C.H., Haring D.R. and W.-S. T.G. Systhesis of counter with threshold elements. In *Record of 6th Annual Symposium on Switching Circuit Theory and Logical Desing*, pages 25–35. IEEE conferance, 1965.

[2] M. Sato and C. Tanaka. Characterization and constructions of reverberating networks. *Mathematical Biosciences*, 62:201–217, 1982.