Additive neural networks and periodic patterns

Tomáš Gedeon*
Department of Mathematical Sciences
Montana State University
Bozeman, MT 59717-0240
gedeon@poincare.math.montana.edu

April 8, 1998

*Research was partially supported by MONTS grant number 219627 and NSF grant DMS-291222

Additive neural networks and periodic patterns

Abstract

In this contribution we discuss weight selection which allows additive neural networks to represent certain periodic patterns. Given a periodic set of vectors V_l whose components are $v_i^l = \pm 1$ we measure correlation between i-th and j-th components of V_l in time l. We show that in the additive neural net with weights chosen based on this correlation, almost all trajectories converge to a periodic orbit, which consecutively visit orthants, determined by the vectors V_l.

We also construct two weights selection processes, one discrete in time and one continuous in time, which construct the desired weights dynamically.

Key Words: Additive neural networks, periodic orbits, weight selection, cascade of nets.

1 Introduction

In this paper we will discuss recurrent additive networks with continuous time activation dynamics

$$\dot{x}_i = -a_i x_i + \sum_{j \neq i} w_{ij} g_j(x_j), \qquad i = 1, \ldots, n, \tag{1}$$

where $n \geq 3$, $a_i > 0$ for all i, $\frac{dg_j}{dx_j} > 0$ for all j, and functions g_j are C^1 and bounded.

Systems of this type are used as models of short term memory, global pattern formation and content-addressable memory. The study of this model and this interpretation goes back to Grossberg [3, 4] and Hopfield [9].

In an application of system (1) to a problem of content addressable memory one interprets equilibria of the activation dynamics as a stored memory and the activation dynamics itself as the retrieval of this memory. In this interpretation, given a memory vector(s) to be stored in the system, the goal is to find weights w_{ij} such that the system (1) admits asymptotically stable equilibrium (equilibria) with coordinates given by memory vector(s). Usually the memory vectors consists of 0's and 1's (or, alternatively, ± 1), and so the desired equilibria are some of the corners of the hypercube in $\mathbf{R}^n$.

In this paper we investigate the ability of additive neural networks (1) to represent time dependent patterns. We make two simplifying assumptions. We shall consider only periodic patterns and each pattern will be represented as a discrete set of vectors, rather then a continuous, periodic function of time.

In the first part of the paper we propose a generalization of Hopfield's [8] weight selection, which assigns a weight w_{ij} based on a correlation in time between i-th and j-th component of the periodic set of vectors.

In the second part of the paper we discuss how the weights can be chosen dynamically. We can think of a periodic pattern in two different ways. We may know the whole set of vectors which form a periodic pattern in advance, or, we can be presented these vectors one at a time. In the first case we can base the construction of the weights on the knowledge of the whole pattern, while in the second case we construct weights from the partial information which is available at that time.

We construct a continuous time weight selection mechanism which is based on the complete knowledge of the pattern and a discrete time weight selection mechanism based on the partial information.

We first define this class of patterns and then the choice of weights. A *periodic pattern* is biinfinite sequence of n-vectors $\{V_l\}_{l=-\infty}^{\infty}$, $n \geq 3$, where each vector $V_l = (v_l^1, \ldots, v_l^n)$ has components $v_l^i \in \{\pm 1\}$, and there exists a permutation
$\alpha \ : \ \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ such that

$$v_l^k v_{l-1}^k = \begin{cases} -1 & \text{if } \ k = \alpha\,(l \ mod \ n) \\ 1 & \text{otherwise} \end{cases} \ . \tag{2}$$

This definition implies that two consecutive states V_l and V_{l-1} of a periodic pattern differ only in a single $\alpha(l \ mod \ n)$ component. Since α is a permutation, every component has to change exactly once in the set of states $V_l, \ldots V_{l+n}$ for any l. It follows that the sequence $\{V_l\}_{l=-\infty}^{\infty}$ is $2n$ periodic. We shall also use the notation $\{V_l\}_{l=1}^{2n}$.

Example. The set of vectors

$$\{V_l\}_{l=1}^{6} = \{(1,1,1), (1,-1,1), (-1,-1,1), (-1,-1,-1), (-1,1,-1), (1,1,-1)\}$$

is a periodic pattern with the permutation $\alpha = (1,2,3) \rightarrow (2,1,3)$.

We denote the open orthants in $\mathbf{R}^n$

$$\mathcal{O}(\sigma_1, \ldots, \sigma_n) := \{x \in \mathbf{R}^n \mid \sigma_i x_i > 0\}$$

where $\sigma_i \in \{\pm 1\}$. The set $(\sigma_1, \ldots, \sigma_n)$ is the *signature* of the orthant $\mathcal{O}(\sigma_1, \ldots, \sigma_n)$. We say that *almost all trajectories converge to a set* K if for every bounded set B containing K the following holds: the set of all $x \in B$ such that the trajectory starting in x converges to K, has the measure of the set B.

Definition 1.1 *A periodic pattern* $\{V_l\}_{l=1}^{2n}$ *is represented* by (1) with a set of weights $W = (w_{ij})$ if almost all trajectories under (1) converge to a trajectory, which consecutively visits orthants

$$\mathcal{O}(V_l), \qquad l = 1, \ldots, 2n.$$

Let $G := \Pi_{i=1}^n g_i'(0)$ be product of gains in system (1). The main result of this paper is the following Theorem.

Theorem 1.2 *Let $\{V_l\}_{l=1}^{2n}$ be a periodic pattern, where each V_l is a n-vector with $n \geq 3$. Let*

$$C_{ij} := \frac{1}{2n} \sum_{l=1}^{2n} v_l^i v_{l-1}^j$$

and

$$w_{ij}^* = \begin{cases} C_{ij} & \text{if } |C_{ij}| = 1 \\ 0 & \text{if } |C_{ij}| < 1 \end{cases}. \tag{3}$$

Then the pattern $\{V_j\}_{j=1}^{2n}$ is represented by the system

$$\dot{x}_i = -a_i x_i + \sum_{j \neq i} w_{ij}^* g_j(x_j) \tag{4}$$

provided that the functions $g_j(x_j)$ are such that $G > 8$ and the origin is hyperbolic. Thus almost all trajectories of (4) converge to an orbit which consecutively visits orthants $\mathcal{O}(V_l), l = 1, \ldots, 2n$.

The numbers C_{ij} measure average correlation (in time) of $i-$th and $j-$th component of the input pattern; the weights w_{ij}^* reflect this correlation.

We want to compare this weight selection to the weight selection of Hopfield [8] for time-independent patterns. Given a set of constant vectors V_s, $s = 1, \ldots, l$ to be stored in a discrete time analog of (1), Hopfield proposed that

$$w_{ij} = \sum_s (2v_i^s - 1)(2v_j^s - 1). \tag{5}$$

Hopfield assumes that $V_s = (v_1^s, \ldots, v_n^s)$ and $v_i^s \in \{0, 1\}$ and so this weight selection measures correlation of i-th and j-th entry over the ensemble of vectors V_s. In [9] Hopfield studies continuous time model (1) with functions $g_j(x_j) = \sigma_j(\kappa x_j)$, where where σ is a sigmoidal function and κ controls the steepness (gain) of σ in such a way that as $\kappa \to \infty$, σ_j approaches a step function. He outlines the proof, which was corrected and completed by Troyer [17], that if all the self-weights w_{ii} are strictly positive, then for high enough gain of functions g_j, at the stable equilibrium point p of (1) the output are saturated. In other words, as $\kappa \to \infty$, the value $g_j(p_j)$ is arbitrary close to one of the limiting values of σ_j at $\pm\infty$. This shows that for sufficiently large gain, the long time behavior of the continuous time system is similar to the long time behavior of the discrete time system.

Our weight selection generalizes the choice of weights (5) since if $V_l = V_k$ for all l, k, (3) gives $|C_{ij}| = 1$ for all ij and we recover the weight selection (5). The only difference is that our vectors consist of ± 1 and not of zeroes and ones.

We prove Theorem 1.2 in section 2. In section 3 we propose weight adjustment mechanisms (one discrete and one continuous in time) which will generate the weights w_{ij}^* dynamically.

The discrete time mechanism is based on Hebb's paradigm [6] according to which strengths of connections are adjusted proportionally to the correlation between the firing patterns of the connected neurons.

In the continuous time weights adjustment mechanism we use the periodic pattern $\{V_l\}$ to construct the set of initial weights, which then evolves toward the weights w^*_{ij}. We use the results for *cascade of neural nets* [7], [16]) to show that the system (1), combined with the weights selection dynamics, will represent the desired periodic pattern.

2 Realization of periodic patterns

In this section we prove Theorem 1.2, while delegating more technical part of the argument to section 4.

Recall that every $2n$-periodic pattern $\{V_l\}_{l=1}^{2n}$ is characterized by a permutation $\alpha(l)$.

Lemma 2.1 *The quantity $|C_{ij}| = 1$ if and only if $(i,j) = (\alpha(s), \alpha(s-1))$ for some $s \in \{1,\ldots,n\}$.*

Furthermore, if $|C_{ij}| \neq 1$ then $1 - |C_{ij}| \geq 2/n$.

Proof. Fix the pair (i,j) and define $a_l := v^i_l v^j_{l-1}$. With this definition $C_{ij} = \frac{1}{2n}\sum_{l=1}^{2n} a_l$. It follows that $|C_{ij}| = 1$ if and only if $a_l a_{l-1} = 1$ for all l. Observe that

$$\begin{aligned} a_l a_{l-1} &= v^i_l v^j_{l-1} v^i_{l-1} v^j_{l-2} \\ &= v^i_l v^i_{l-1} \cdot v^j_{l-1} v^j_{l-2}. \end{aligned}$$

Recall that by the definition of the periodic pattern $v^i_l v^i_{l-1} = -1$ if $i = \alpha(l)$, otherwise $v^i_l v^i_{l-1} = 1$. Therefore $a_l a_{l-1} = 1$ if and only if both

$$i \neq \alpha(l) \quad \text{and} \quad j \neq \alpha(l-1).$$

However, as l ranges from 1 to $2n$, there must be a value of l, say $l = s$, such that

$$i = \alpha(s). \tag{6}$$

Since the pattern is $2n$ periodic, we also have, in addition to (6), that $i = \alpha((s+n)\, mod\, n)$.

Now there are two possibilities. If $j = \alpha(s-1)$, then $a_s = -1$ and $a_{s-1} = -1$ and, consequently, $a_l = a_{l-1}$ for all l. In this case $|C_{ij}| = 1$.

If $j \neq \alpha(s-1)$ then $a_s = -1$ and $a_{s-1} = 1$, and, by the periodicity, $a_{s+n} = -1$ and $a_{s-1+n} = 1$. In this case it is easy to see that $1 - |C_{ij}| \geq 2/n$. □

Let us consider system (4). In the light of Lemma 2.1 which weights will be non-zero depends on the permutation α, which characterizes a simple pattern $\{V_l\}_{l=1}^{2n}$.

In order to simplify discussion, we introduce a change of variables, which puts the system (4) to a standard form. Given a periodic pattern $\{V^\alpha_l\}_{l=1}^{2n}$ in $\mathbf{R}^n$ with a permutation α, we define a change of coordinates

$$v_i \rightarrow v_{\alpha(i)}. \tag{7}$$

Observe that in the new coordinates the periodic pattern becomes $\{V_l^{id}\}_{l=1}^{2n}$, where id is the identity permutation. It is straightforward to check that for such a periodic pattern the limiting weights have the form

$$w_{ij}^* = \begin{cases} -1 & \text{if } i = 1, j = n \\ 1 & \text{if } i \neq 1, j = i-1 \\ 0 & \text{otherwise} . \end{cases} \quad . \tag{8}$$

With these weights the system (4) becomes

$$\dot{x}_i = -a_i x_i + w_{i,i-1}^* g_{i-1}(x_{i-1}), \qquad i = 1, \ldots, n, \tag{9}$$

where $w_{1,n}^* = -1$ and all other $w_{i,i-1}^* = 1$.

We call the pattern $\{V_l^{id}\}_{l=1}^{2n}$ the *standard form of a periodic pattern* and (9) is *the standard form* of the equation (4).

Theorem 2.2 *Assume that $G > 8$ and that the origin in (9) is hyperbolic. Then the only asymptotically stable invariant sets in (9) are periodic orbits, which visit consecutively $2n$ orthants $\{\mathcal{O}(V_l^{id})\}_{l=1}^{2n}$ with signatures*

$$\begin{aligned} \{V_l^{id})\}_{l=1}^{2n} \;=\; & \{(1,1,\ldots,1,1),(-1,1,\ldots,1,1),(-1,-1,\ldots,1,1),\ldots,(-1,-1,\ldots,-1,-1), \\ & (1,-1,\ldots,-1,-1),\ldots,(1,1,\ldots,1,-1)\}. \end{aligned} \tag{10}$$

Furthermore, almost all trajectories converge to such a periodic orbit.

If we reverse the change of variables (7) Theorem 2.2 implies Theorem 1.2.

Proof of the Theorem 2.2 is delegated to section 4.

3 Weight selection mechanisms

In this section we present two weight selection mechanisms which compute weights w_{ij}^*.

The first system is discrete time system and the pattern $\{V_l\}$ affects the dynamics of the system as a non-autonomous input. The adjustment of the weights follows Hebbian paradigm [6]: the strength of the connection w_{ij} is proportional to the correlation of the activity of the neurons i and j.

The second system is a continuous time dynamical system. In this case the periodic pattern $\{V_l\}$ gives rise to an initial condition of the weight selection dynamics.

Neither the discrete nor the continuous weight selection is flexible in a sense that upon presentation of a new periodic pattern, it will select the weights for the representation of a new pattern without outside adjustment. In both cases the weight selection process must be restarted.

3.1 Discrete weight selection dynamics

Fix a pair i, j. Let

$$\beta_{ij}(s) := v_i^{s+1} v_j^s \quad \text{for all } s.$$

The discrete weight adjustment dynamics is

$$w_{ij}^s = \frac{1}{2} w_{ij}^{s-1} + \frac{U_{ij}(s)}{2} \tag{11}$$

where $U_{ij}(s)$ is defined inductively by $U_{ij}(1) := \beta_{ij}(1)$ and

$$U_{ij}(s) := \frac{1}{2}\beta_{ij}(s) U_{ij}(s-1)(U_{ij}(s-1)) + \beta_{ij}(s)).$$

The number s represents the discrete time and the initial condition w_{ij}^0 for w_{ij} is arbitrary.

Proposition 3.1 *The iterates of the non-autonomous map (11) converge to the weights (3), i.e. for every pair (i,j)*

$$w_{ij}^s \to w_{ij}^*. \qquad \textit{as } s \to \infty$$

Proof. We fix a pair i, j and drop the subscript from $U_{ij}(s)$ and $\beta_{ij}(s)$.

Since $\beta(s) \in \{\pm 1\}$, the value of $U(s)$ can only be $0, 1$ or -1. We make some simple observations. If $U(l-1)\beta(l) = -1$ then $U(l-1) + \beta(l) = 0$ and so $U(l) = 0$. If $U(l) = 0$ then $U(s) = 0$ for all $s \geq l$. It follows that if $U(1) = 1$ then either $U(s) = 1$ for all s, or there is K such that $U(s) = 1$ for $s = 1, \ldots, K$ and $U(s)) = 0$ for all $s > K$. Similar statement is true for the case $U(1) = -1$.

Function $U(s)$ serves as a memory; if the last entry $U(s-1)$ matches the input $\beta(s)$ then $U(s) = U(s-1)$, in the opposite case, $U(s) = 0$. It is easy to see that if the periodic pattern has period $2n$, then $U(s) = U(n)$ for all $s \geq n$.

Thus for $s \geq n$ equation (11) has three possible forms corresponding to the three possible values of $U(s) \in \{\pm 1, 0\}$. It is easy to see from the form of (11) that if $U(s) = \pm 1$ for all $s \geq n$ then $w_{ij}^s \to \pm 1$ as $s \to \infty$, and if $U(s) = 0$ for all $s \geq n$, then $w_{ij}^s \to 0$ as $s \to \infty$.

Observe that if $U_{ij}(s) = 1$ for all s we must have that $\beta_{ij}(s) = 1$ for all s and thus $C_{ij} = 1$. By selection rule (3), $w_{ij}^* = 1$ which means that $w_{ij}^s \to w_{ij}^*$ as $s \to \infty$. Similarly, if $U_{ij}(s) = -1$ for all s, then $C_{ij} = -1$ and $w_{ij}^s \to w_{ij}^*$ as $s \to \infty$.

If $U_{ij}(s) = 0$ for $s \geq k$ we must have that $\beta_{ij}(l)\beta_{ij}(l-1) < 0$ for some $l \leq k$ and so $|C_{ij}| < 1$. We see that also in this case $w_{ij}^s \to w_{ij}^* = 0$ as $s \to \infty$.

This proves the Proposition. □

Remark 3.2 To restart the weight selection process for a new pattern, one needs to reset the function $U_{ij}(1) = \beta_{ij}(1)$ for all pairs i, j.

3.2 Continuous weight selection dynamics

Given a periodic pattern $\{V_l\}_{l=1}^{2n}$ we define for every pair $i, j \in \{1, \ldots, n\} \times \{1, \ldots, n\}$ a set of $2n$ variables w_{ij}^l, $l = 1, \ldots, 2n$. A periodic pattern $\{V_l\}_{l=1}^{2n}$ gives rise to a set of initial data

$$w_{ij}^l(0) := v_l^i v_{l-1}^j. \tag{12}$$

Let $W := (w_{11}^1, w_{11}^2, \dots, w_{nn}^{2n}) \in \mathbf{R}^{2n^3}$ be the vector of weights and let $W(0)$ be the vector of initial weights (12).

The weight adjustment dynamics will be given by

$$\dot{w}_{ij}^l = w_{ij}^{l-1} - 2w_{ij}^l + w_{ij}^{l+1} + \epsilon f(w_{ij}^l) \tag{13}$$

where $l = 1, \dots, 2n$ is taken mod $2n$, the subscript $ij \in \{1, \dots, n\} \times \{1, \dots, n\}$ and

$$f(u) = -u(u^2 - 1)(u^2 - (1 - \frac{1}{2n})^2). \tag{14}$$

The number ϵ is small and will be specified later.

Recall that a dynamical system is *dissipative* if there is a bounded set B such that all trajectories eventually enter this set. Since for $|y| >> 1$ we have $f(y)y < 0$ it is easy to see that the system (3) is dissipative.

Observe that the system (13) involves $2n^3$ differential equations; however equations involving w_{ij}^l, $l = 1, \dots, 2n$, and w_{qs}^m, $m = 1, \dots, 2n$ for $(i,j) \neq (q,s)$ are decoupled. So (13) is composed of n^2 subsystems of $2n$ equations each, which are mutually independent. We shall call a subsystem involving w_{ij}^l, $l = 1, \dots, 2n$ the (i,j)*-subsystem*. Each subsystem will determine one weight in (4).

Following Proposition is an analog of Proposition 3.1 for continuous time weight selection.

Proposition 3.3 *There is ϵ_0 such that for all $\epsilon \leq \epsilon_0$, the trajectory in the system (13) with the initial condition (12) converges to an asymptotically stable equilibrium*

$$W^* := (w_{11}^1, w_{11}^2, \dots, w_{11}^{2k}, w_{12}^1, \dots, w_{nn}^{2n})$$

where

$$w_{ij}^l = w_{ij}^m = w_{ij}^*$$

for all i, j, l, m.

To prove Proposition 3.3 we need a few Lemmas.

We first consider the system (13) with $\epsilon = 0$:

$$\dot{w}_{ij}^l = w_{ij}^{l-1} - 2w_{ij}^l + w_{ij}^{l+1}. \tag{15}$$

Observe that the system (15) represents a discrete diffusion process. Following considerations are standard and we include them for the sake of completeness. Let

$$A_{ij}(t) := \frac{1}{2n} \sum_{l=1}^{2n} w_{ij}^l(t)$$

be the average of the variables w_{ij}^l, $l = 1, \dots, 2n$, at time t under the flow given by (15).

Lemma 3.4 *Under the flow given by (15), $A_{ij}(t) = A_{ij}(0)$ for all t.*

Proof.

$$\begin{aligned}\frac{d}{dt}A_{ij}(t) &= \frac{1}{2n}\sum_{l=1}^{2n}\dot{w}_{ij}^l(t)\\ &= \frac{1}{2n}\sum_{l=1}^{2n} w_{ij}^{l-1} - 2w_{ij}^l + w_{ij}^{l+1} = 0,\end{aligned}$$

since l is taken mod $2n$. $\square$

Lemma 3.5 *Under the flow given by (15), $w_{ij}^l(t)$ converges to an equilibrium $\bar{w}_{ij}(l)$ for any i,j and l. Furthermore, $\bar{w}_{ij}^l = A_{ij}$ is independent of l.*

Proof. Consider the function

$$V_{ij}(t) := \frac{1}{2}\sum_{l=1}^{2n}(w_{ij}^l(t) - A_{ij})^2.$$

We show that this is a Lyapunov function for every pair i, j.

$$\begin{aligned}\dot{V}_{ij}(t) &= \sum_{l=1}^{2n}(w_{ij}^l - A_{ij})\dot{w}_{ij}^l\\ &= \sum_{l=1}^{2n}(w_{ij}^l - A_{ij})(w_{ij}^{l-1} - 2w_{ij}^l + w_{ij}^{l+1})\\ &= \sum_{l=1}^{2n} w_{ij}^l(w_{ij}^{l-1} - 2w_{ij}^l + w_{ij}^{l+1})\\ &= \sum_{l=1}^{2n} w_{ij}^l w_{ij}^{l-1} - 2(w_{ij}^l)^2 + w_{ij}^l w_{ij}^{l+1}\\ &= -\sum_{l=1}^{2n}(w_{ij}^l)^2 - 2w_{ij}^l w_{ij}^{l-1} + (w_{ij}^{l-1})^2\\ &= -\sum_{l=1}^{2n}(w_{ij}^l - w_{ij}^{l-1})^2\end{aligned}$$

where we used repeatedly the fact that l is taken mod $2n$. We see that $\dot{V}_{ij}(t) \le 0$ for all t. Further, $\dot{V}_{ij}(t) = 0$ if and only if $w_{ij}^l = w_{ij}^{l-1}$ for all l i.e. when $w_{ij}^l = w_{ij}^m$ for all l, m. It follows from Lemma 3.4 that this common value is A_{ij}. $\square$

For every (i,j)-subsystem we denote by $R_{ij} := \{w_{ij}^l \mid -1 \le w_{ij}^l \le 1\}$ the cube centered in the origin of $\mathbf{R}^{2n}$. We observe that the initial vector of (i,j)-weights $w_{ij}(0) = (w_{ij}^1(0), w_{ij}^2(0), \ldots, w_{ij}^{2n}(0))$ is one of the corners of this cube, since $w_{ij}^l \in \{\pm 1\}$. Let us denote $\mathbf{1} := (1, \ldots, 1) \in \mathbf{R}^{2n}$, $-\mathbf{1} := (-1, \ldots, -1) \in \mathbf{R}^{2n}$ and $\mathbf{0} := (0, \ldots, 0) \in \mathbf{R}^{2n}$.

Let J be the set of corners in R_{ij} with the following property. If $w \in J$, $w = (w^1, \ldots, w^{2n})$, then

$$-1 + 2/n \le A_w \le 1 - 2/n$$

where $A_w := \frac{1}{2n}\sum_{i=1}^{2n} w^i$.

Lemma 3.6 *Fix a pair (i,j). There exists an ϵ_0 such that for all $\epsilon \leq \epsilon_0$ every trajectory of (13) with initial value in J converges to* $\mathbf{0}$.

Proof. Assume that $w(0) \in J$ and so $A_w(0) \in [-1+2/n, 1-2/n]$. It follows from Lemma 3.5 that there is a time T such that $w(t) \in [-1+1/n, 1-1/n]$ for all $t \geq T$ under the flow (15). By the continuity of the flow there is $\epsilon > 0$ such that

$$w(t) \in [-1+1/n, 1-1/n]$$

for $t \geq T$ under the flow (13).

Let

$$M(t) := \max_{i=1,\ldots,2n} \quad w^i(t) \quad \text{and} \quad m(t) := \min_{i=1,\ldots,2n} \quad w^i(t).$$

Let $w^l(t)$ is an element maximizing $M(t)$ at time t. Then

$$\dot{M}(t) = \dot{w}^l(t) = w^{l+1} - 2w^l + w^{l-1} + \epsilon f(w^l) < 0$$

since $f(w^l) < 0$ and $w^{l+1} - 2w^l(t) + w^{l-1} \leq 0$ by the fact that $w^l(t)$ maximizes $M(t)$. Thus $M(t)$ is a monotone decreasing function. Analogous argument shows that $m(t)$ is a monotone increasing function.

It follows that all $w^l(t)$ converge to the same value and that this value is in the interval $[-1+1/n, 1-1/n]$. It follows from the form of the function $f(x)$ that $w^l(t) \to 0$ for all l and so $w(t) \to \mathbf{0}$ as $t \to \infty$. $\square$

Proof of Proposition 3.3 Let $w_{ij}(0)$ be the initial weight vector for (i,j) subsystem. Observe that $A_{ij}(0) = C_{ij}$ where the later quantity is used in (3) to select weights w^*_{ij}. From Lemma 2.1 we get that $|A_{ij}(0)| = 1$ if and only if $i, j \in I_l$ and $i = \alpha_l(s), j = \alpha_l(s-1)$ for some s and where α_l is the permutation for simple periodic pattern on I_l. If $|A_{ij}(0)| \neq 1$ then again by Lemma 2.1 we have that $1 - |A_{ij}(0)| \geq 2/k$. In other words, $w_{ij}(0) \in J$.

Thus, for each pair i,j either $w_{ij}(0)$ is such that $|A_{ij}(0)| = |C_{ij}| = 1$, in which case

$$w_{ij}(t) \to C_{ij} = w^*_{ij}$$

or $w_{ij}(0) \in J$, in which case by Lemma 3.6

$$w_{ij}(t) \to 0 = w^*_{ij}.$$

Since there are finitely many (i,j)- subsystems where $w_{ij}(0) \in J$, there is an $\epsilon > 0$ such that all subsystems converge for $\epsilon \leq \epsilon_0$.

It remains to be shown that the vector W^* is an asymptotically stable equilibrium. It is enough to show that the vectors $\mathbf{1}$, $-\mathbf{1}$ and $\mathbf{0}$ are asymptotically stable equilibria in any (i,j) subsystem. Observe that for small ϵ the linearization of (15) in each of these equilibria has $2n-1$ eigenvalues with negative real part and a zero eigenvalue with the eigenvector $(1,1,\ldots,1)$. For ϵ small enough the linearization of (13) still has $2n-1$ eigenvalues with negative real part. However, from the form of the equation (13) follows that the last eigenvalue is $\epsilon f'(q)$ for $q = 0, \pm 1$. Since $f'(q) < 0$ for $q = 0, \pm 1$ the vectors $\mathbf{1}$, $-\mathbf{1}$ and $\mathbf{0}$ are asymptotically stable equilibria in any (i,j) subsystem.

$\square$

3.3 Real time weight selection

We now turn to the question of real time weight adjustment. We consider weight adjustment system (13) and the activation system (1) together and study the dynamics of the combined system. Recall, that the system (13) contains $2n$ copies w_{ij}^l, $l = 1, \ldots, 2n$ of the weight w_{ij}; we need to chose which weight will appear in the system (1). In the light of Theorem 3.3 this choice does not matter, since all w_{ij}^l, $l = 1, \ldots, 2n$, converge to the same value w_{ij}^*. We let $w_{ij} := w_{ij}^1$ for all pairs i, j. Consider a *cascade of neural nets* ([7], [16])

$$\begin{aligned} \dot{w}_{ij}^l &= w_{ij}^{l-1} - 2w_{ij}^l + w_{ij}^{l+1} + \epsilon f(w_{ij}^l) \\ \dot{x}_i &= -a_i x_i + \sum_{j \neq i} w_{ij}^1 g_j(x_j), \end{aligned} \tag{16}$$

where $i, j \in \{1, \ldots, n\}, l \in \{1, \ldots, 2n\}$.

We studied both systems separately; the weight adjustment system converges to an asymptotically stable equilibrium W^* by Theorem 3.3 and the activation dynamics with weights fixed at the value given by W^* is considered in Theorem 1.2.

Theorem 3.7 *Consider a periodic pattern $\{V_l\}_{l=1}^{2n}$ of n-vectors, $n \geq 3$. Assume that the system*

$$\dot{x}_i = -a_i x_i + \sum_{j \neq i} w_{ij}^* g_j(x_j)$$

with weights $W^* = (w_{ij}^*)$ satisfy the assumption of Theorem 1.2 and that $f(x)$ is as in (14).

Then for almost all vectors $\mathbf{x} = (\mathbf{x_1}, \ldots, \mathbf{x_n}) \in \mathbf{R}^n$ the system (16) with initial data $(\mathbf{x}, W(0))$, where the initial weights $W(0)$ are given by (12), converges to a periodic orbit, which generates periodic pattern $\{V_l\}_{l=1}^{2n}$.

The proof is postponed to section 5.

4 Convergence to a periodic orbit

We shall prove all results in this section for the equation in the standard form (9). One can get the results for the general equation (4) by reversing the change of coordinates (7).

The system (9) is a *cyclic feedback system* which was studied in [2], [10] and [13]. Early results were given in [5]. Since the nonlinearities g_i are monotone this is a *monotone cyclic feedback system.* The cyclic feedback systems fall into two large categories; those with *positive* and those with *negative* feedback. Since the product $\prod_{i=1}^n w_{i,i-1}^* = -1$ is negative, (9) is a negative feedback system.

Theorem 4.1 (Mallet-Paret and Smith, [13]) *Consider a monotone cyclic feedback system in $\mathbf{R}^n$. For any bounded trajectory $x(t)$, its omega limit set $\omega(x)$ is one of the following:*

i) a fixed point

ii) a limit cycle

iii) a set $H = E \cup C$ where E is set of equilibria and C is the set of connecting orbits between the equilibria in E.

Furthermore, for any i the projection $\pi_i : \mathbf{R}^n \to \mathbf{R}^2$, given by $(x_1, \ldots, x_n) \to (x_i, x_{i-1})$, is injective on the invariant set.

Last part of the Theorem 4.1 implies that the following definition makes sense.

We say that invariant set S *surrounds* invariant set $\bar{S}$ if for all i we have $\pi_i(\bar{S}) \subset \pi_i(S)$. We shall use this definition later in Theorem 4.5.

Theorem 4.1 is proved using a *discrete Lyapunov function* . The construction of this function can be found in Mallet-Paret and Smith [13]. We recall main points of this construction in our notation.

Definition 4.2 Let $x(t) = \bar{x}(t) - \tilde{x}(t)$ or $x(t) = \dot{\bar{x}}(t)$ for any two solutions $\bar{x}(t)$, $\tilde{x}(t)$ of the system (9). Define

$$N(x) = card\{i \mid w_{i,i-1} x_i x_{i-1} < 0\}$$

if $x_i \neq 0$ for all i (here we use convention $x_1 = x_{n+1}$). We can extend the domain of definition of N by continuity to

$$\mathcal{N} = \{x \in R^n \mid x_i = 0 \text{ implies } w_{i+1,i} w_{i,i-1} x_{i+1} x_{i-1} < 0\}$$

on which N is continuous. If $x \in R^n \setminus \mathcal{N}$ we leave N undefined.

Observe also, that for those $x \in R^n$ with each $x_i \neq 0$, $1 \leq i \leq n$

$$(-1)^{N(z)} = sign \prod_{i=1}^{n} w_{i,i-1} x_i x_{i-1} = \prod_{i=1}^{n} w_{i,i-1} = -1 \tag{17}$$

and so N takes only odd values.

A geometrical view of $\mathcal{N}$ may be enlightening. We observe that $\cup \mathcal{O} \subset \mathcal{N}$ and on each orthant the value of N is constant. Let

$$X_i := \{x \in \mathbf{R}^n \mid x_i = 0, \ w_{i+1,i} w_{i,i-1} x_{i+1} x_{i-1} < 0\}$$

denote those parts of coordinate hyperplanes, which are the boundaries of two open orthants on which N has the same value. Then

$$\mathcal{N} = (\cup X_i) \cup (\cup \mathcal{O}).$$

One can use discrete Lyapunov function N to define a Morse decomposition of the attractor.

A *Morse decomposition* $\mathcal{M}(\mathcal{A}) = \{M(p) \mid p \in (\mathcal{P}, \geq)\}$ of an invariant set $\mathcal{A}$ is a decomposition of $\mathcal{A}$ into at most a finite number of disjoint compact invariant subsets $M(p)$, called *Morse sets*, indexed by a partially ordered set $(\mathcal{P}, \geq)$, such that

1. given $x \in \mathcal{A}$ if $\omega(x) \in M(p)$ and $\alpha(x) \in M(q)$ then $q \geq p$
2. if $\omega(x) \in M(p)$ and $\alpha(x) \in M(p)$ then $\varphi(x,t) \in M(p)$ for all t, where $\varphi : \mathcal{A} \times R \to \mathcal{A}$ denotes the flow.

Using the Lyapunov function, described above, we can define a Morse decomposition $\mathcal{M}(\mathcal{A}) = \{M(p) \mid p = 0, 1, \ldots, P\}$ of the invariant set of (9). Let

$$\tilde{M}(p) := \{x \in R^n \mid N(\varphi(x,t)) = 2p + 1 \text{ for all } t \in R\}.$$

This definition has a geometric interpretation: $\tilde{M}(p)$ is a maximal invariant set which lies in the union of open orthants on which the function N assumes constant value k. Notice that $\{0\}$ is also an invariant set and as such it must be included in some Morse set. However, since $N(0)$ is not defined it is not included in any set $\tilde{M}(p)$ and so these sets do not form a Morse decomposition. We want to add $\{0\}$ to one of the sets $\tilde{M}(p)$ and construct a Morse decomposition. Let us consider x such that $\alpha(x) = \{0\}$ and $\omega(x) = \mathcal{S} \subset \tilde{M}(r)$. By the definition of the Morse decomposition we have to include 0 to some set $\tilde{M}(q)$ with $q \geq r$. Since x is an arbitrary element from the unstable manifold of the origin $W^u(0)$, we must have $q \geq p$ for every p such that there is x with $\alpha(x) = \{0\}$ and $\omega(x) \subset \tilde{M}(p)$. Similar argument applies to the stable manifold $W^s(0)$ of the origin: if $x \in W^s(0)$ and $\alpha(x) \subset \tilde{M}(s)$ then $s \geq q$. It can be shown by analyzing the linearization of (9) at the origin ([13]), that $N(W^s(0) \cap B_\epsilon) > N(W^u(0) \cap B_\epsilon)$ for a small ball B_ϵ around the origin. Since N is nonincreasing along the trajectories this rules out the existence of a homoclinic orbit to the origin. Therefore we can lump together the set $\{0\}$ with the sets $\tilde{M}(p)$ "above" $\{0\}$ into one Morse set and let the sets $\tilde{M}(p)$ "below" $\{0\}$ be the other Morse sets. Thus we define

$$\begin{aligned} M(p) &:= \tilde{M}(p) \qquad p < P \\ M(P) &= \{0\} \cup \{\cup_{p \geq P} \tilde{M}(p)\}. \end{aligned}$$

The value of P, not surprisingly, will depend on the dimension J of the unstable manifold of the origin ([10]): If $J = 2i, 2i+1$ then $P = i - 1$ and if $J = n$ then $P = n + 1/2$ if n is odd and $P = n/2$ if n is even.

Lemma 4.3 *For the system (9) we have $J \geq 2$ and so $P > 0$.*

Proof. It is straightforward to compute the characteristic polynomial of the linearization of (9) at zero. We get

$$(\lambda + 1)^n = \Pi_{i=1}^n w_{i,i-1} g_i'(0) = \Pi_{i=1}^n w_{i,i-1} \Pi_{i=1}^n g_i'(0) = -G$$

by definition of gain G. It is easy to check that the assumptions $n \geq 3$ and $G > 8$ imply that there are at least two eigenvalues in the right half plane. Thus $J \geq 2$ and $P > 0$ using the formulas above. □

Remark 4.4 Note that this is the only place where we use the assumption $G > 8$.

The Morse decomposition exhibits the gradient-like properties of the flow on $\mathcal{A}$ and confines all recurrent dynamics into individual Morse sets. It also gives a rough idea about the stability of various invariant sets; the most stable sets should be in the lowest Morse set i.e. the set on the bottom of the partial order. However, one must be cautioned that the asymptotic stability of an invariant set S is in general not related to the Morse ordering of the Morse set, in which S lies. In fact there may be an asymptotically stable invariant set in other then the lowest Morse set. This was shown for the Morse decomposition of the attractor of the scalar delay equation with negative feedback by Ivanov and Losson [11].

However, for the monotone cyclic feedback systems J. Mallet-Paret [14] recently proved a remarkable theorem:

Theorem 4.5 (Theorem 1, [14]) *Consider a monotone cyclic feedback system. Let $q(t)$ be a non-constant periodic solution, $E \in \mathbf{R}^n$ be an equilibrium and assume that solution $q(t)$ surrounds the equilibrium E. Let $\nu = N(q(t) - E)$, where N is the discrete Lyapunov function described above. Then the dimension of the center-unstable manifold $W^{cu}(q(\cdot))$ of the periodic solution $q(\cdot)$ satisfies*

$$\nu \leq \dim W^{cu}(q(\cdot)) \leq \nu + 1.$$

Consider a periodic orbit q and let $\sigma_1 \leq \sigma_2 \leq \ldots \sigma_n$ be norms of Floquet multipliers for the linearization of (9) about q. Let F_σ be the corresponding generalized eigenspace of the period map.

Theorem 4.6 (Theorem 2.6, [13]) *The norms of Floquet multilpiers satisfy*

$$\sigma_1 \leq \sigma_2 < \sigma_3 \leq \sigma_4 < \ldots < \sigma_{n-2} \leq \sigma_{n-1} < \sigma_n \tag{18}$$

and the value of N on $F_{\sigma_{2p+1}} + F_{\sigma_{2p+2}}$ is $2p+1$ for $p = 0, 1, \ldots$.

Note that whether the last inequality is strict or not depends on the parity of n; the case of n odd is illustrated in (18).

Corollary 4.7 *For every periodic orbit $q(t) \notin M(0)$ in a monotone negative cyclic feedback system, the dimension of center-unstable manifold $\dim W^{uc}(q) \geq 3$. Furthermore, there are at least two Floquet multipliers with modulus larger than one and so such an orbit is unstable.*

Proof. From the definition of Morse decomposition for negative cyclic feedback system follows that the value of function N on $M(p)$ is $2p+1$. If $q(t) \in M(p)$, $p \geq 1$ then $N(q(t)) = N(q(t) - 0) \geq 3$. Using Theorem 4.5 with $E = 0$ it follows that $dim\ W^{uc}(q) \geq 3$. Using (18) this in turn implies that $\sigma_1 > 1$ and $\sigma_2 > 1$. □

The set $M(0)$ is characterized by the fact that $N(x(t)) = 1$ for all $x \in M(0)$ and all t. This is a geometrical condition saying that $M(0)$ must be a subset of family of orthants, on which $N(x) = 1$. We have the following Proposition.

Proposition 4.8 *Consider system (9). The function $N(x)$ has value 1 on the $2n$ orthants $\mathcal{O}(V_l)$ where $\{V_l\}_{l=1}^{2n}$ is given in (10).*

Proof. We have to show that N has value 1 on all the orthants $\{\mathcal{O}_l(V_l^{id})\}_{l=1}^{2n}$ where $\{V_l^{id}\}_{l=1}^{2n}$ is given in (10).

Let us consider $x \in \mathcal{O}(V_l)$ for some l. We want to show that $N(x) = 1$. By definition

$$N(x) = card\{\#i \mid w^*_{i,i-1} x_i x_{i-1} < 0\} = card\{\#i \mid w^*_{i,i-1} (sign\ x_i)\ (sign\ x_{i-1}) < 0\}.$$

We have $sign\ x_i = v_l^i$ and $sign\ x_{i-1} = v_l^{i-1}$. Furthermore, since the weights $w^*_{i,i-1}$ are nonzero we have $w^*_{i,i-1} = v_k^i v_{s-1}^{s-1}$ for any s, in particular for $s = l$. So

$$N(x) = card\{\#i \mid w_{i,i-1}x_i x_{i-1} < 0\} = card\{\#i \mid (v_l^i)^2 v_l^{i-1} v_{l-1}^{i-1} < 0\}.$$

We observe that if $i-1 = l$ then $v_l^{i-1}v_{l-1}^{i-1} = v_l^l v_{l-1}^l = -1$ (see (2)). For all other i the product $v_{l-1}^{i-1}v_l^{i-1} = 1$. Since $(v_l^i)^2 = 1$ we have that $N(x) = 1$. The choice of the orthant $\mathcal{O}(V_l)$ was arbitrary, and so $N(x) = 1$ on all the orthants $\mathcal{O}(V_l)$. □

It is clear that for the understanding of a global behavior of a cyclic feedback system it is important to know the structure of the lowest Morse set. We will call a periodic orbit $x(t)$ *large* if, for all $i = 1, \dots, n$, there are times t_i and t_i' such that $[x(t_i)]_i > 0, [x(t_i')]_i < 0$, where $[x]_i$ denotes the i-th coordinate of x. Large periodic orbits describe oscillations with the property, that each variable changes sign twice during one period of oscillation.

We shall use following results from Gedeon and Mischaikow [10]:

Theorem 4.9 (Theorem 1.4, [10]) *Consider a negative monotone cyclic feedback system. Let $p = 0, \dots, P-1$. If $M(p)$ contains no fixed points then $M(p)$ contains a large periodic orbit.*

Lemma 4.10 (Lemma 5.1, [10]) *Given a monotone cyclic feedback system, assume that $M(p)$ contains no fixed points. Let $\mathcal{O} = \mathcal{O}(\sigma_1, \dots, \sigma_n)$ be an orthant such that $\mathcal{O} \cap M(p) \neq \emptyset$. Then* $\mathrm{Inv}\mathcal{O} = \emptyset$.

Lemma 4.11 (Corollary 5.3, [10]) *For any recurrent set S of a cyclic feedback system one of the following holds:*

- *$S \subset \mathcal{O}(\sigma_1, \dots, \sigma_n)$ for some collection $\{\sigma_i\}$*
- *S is large.*

Proof of Theorem 1.2

We prove the Theorem for the standard form (9). We first show that there are no fixed points in (9) besides the origin. Indeed, the x_i component of such a fixed point satisfies $\dot{x}_i = 0$ which means $x_i = w^*_{i,i-1}g_{i-1}(x_{i-1})$. Since the same holds for $x_{i-1}, x_{i-2}, \dots, x_{i+1}$ components we get

$$x_i = w^*_{i,i-1}g_{i-1}(w^*_{i-1,i-2}g_{i-2}(\dots w^*_{i+1,i}g_i(x_i))) =: h(x_i).$$

Observe that

$$\frac{dh}{dx_i} = \prod_i w^*_{i,i-1} \prod_i \frac{dg_i}{dx_{i-1}} < 0$$

since $\frac{dg_i}{dx_{i-1}} > 0$ for all i and $\prod_i w_{i,i-1} = -1$. Also $h(0) = 0$ by the assumptions on the functions g_i and so the only solution of $x_i = h(x_i)$ is $x_i = 0$. Since i was arbitrary, the only fixed point is the origin.

It follows from Theorem 4.1 that the Morse sets $M(p)$, $p = 0, \dots, P-1$ consist of periodic orbits and the connecting orbits between them. The set of large periodic orbits is non-empty by Theorem 4.9. By Lemma 4.10 and Lemma 4.11 all periodic orbits are large.

By Corollary 4.7 all periodic orbits not in the Morse set $M(0)$ are unstable. We now show that almost all trajectories converge to a periodic orbit in $M(0)$.

We first show that the system (9) is dissipative. By assumption all functions g_i are bounded and so there is a constant c such that

$$\max_{i=1,\ldots,n;x\in\mathbf{R}} \frac{|g_i(x)|}{a_i} < c.$$

Consider any box in $\mathbf{R}^n$ with side $r > 2c$ centered at the origin. Then on the boundary of the box the term $-a_i x_i$ will dominate the term $w_{i,i-1}g_i(x_{i-1})$ and the vector field points inward. It follows that the vector field generated by (9) is dissipative and the set of initial data whose orbits escape to infinity is empty.

Fix a bounded set B, such that all the invariant set is in the interior of B. Let U be the union of periodic orbits which do not belong to $M(0)$ ant the origin. Observe that Corollary 4.7 implies that all periodic orbits in U have at least two Floquet multipliers with the modulus larger then 1. Also, by Lemma 4.3, the origin has at least two eigenvalues wit positive real part.

For any invariant set q we denote the center-stable manifold of the set q by $W^{cs}(q)$. We show that the set of points

$$Q := \{x(0) \in B \cap W^{cs}(q),\ x(t) {\rightarrow} q \text{ as } t {\rightarrow} \infty \mid q \in U\}$$

has measure zero. It then follows that almost all trajectories in B converge to an invariant set in $M(0)$.

Since Morse sets are disjoint the set $U \cap M(i)$ is disjoint from the set $U \cap M(j)$ for $i \neq j$. Since the origin is hyperbolic it is not a limit point of the set U. It follows that $U \cap M(i)$ is compact for all i.

Consider now an isolated component P of the set $U \cap M(i)$. This means that there exists a neighborhood $N_\epsilon(P)$ such that the maximal invariant set in $N_\epsilon(P)$ is P. Observe that the origin is hyperbolic, and so it itself is an isolated component of $U \cap M(P)$. We first concentrate on periodic orbits and so we assume that P consists of periodic orbits.

Since $U \cap M(i)$ is compact and U is isolated, U is compact. Let q be a periodic orbit in P such that there is a sequence of other periodic orbits converging to q. Then q has at least two multipliers m_1, m_2 with modulus 1. It follows from (18) that in fact it has precisely two multipliers with modulus 1. Observe that $m_1 = 1$ since it corresponds to the trivial eigenvector $\dot{q}$ and so m_2 must be real. This implies that $m_2 = 1$ or $m_2 = -1$. It is easy to see that the second case implies existence of another periodic orbit p in the neighborhood of q, with the period approximately twice the period of q, whose projection to x_1, x_2 plane intersect projection of q. This contradicts Theorem 4.1 and so $m_2 = 1$.

We restrict ourself to a Poincaré section Σ to the set P. From (18) and the standard theory of invariant manifolds ([1]) we have that the fixed point $\tilde{q}$ of the Poincaré map, corresponding to q, has a one dimensional center manifold $W^c(\tilde{q})$, at least two dimensional manifold $W^u(\tilde{q})$ and and codimension at least 2 manifold $W^s(\tilde{q})$. Any invariant set under the Poincaré map in $N_\epsilon(\tilde{q}) \cap \Sigma$ is a subset of $W^c(\tilde{q})$. Locally, $W^c(\tilde{q})$ is a curve tangent to the eigenvector of the Poincaré map corresponding to eigenvalue $m_2 = 1$ at $\tilde{q}$.

If $q \in P$ is isolated, then $W^c(\tilde{q}) = \emptyset$.

Let

$$\mathcal{S}_\epsilon(P) := \{x(0) \in N_\epsilon(P),\ x(t) \to q \text{ as } t \to \infty, q \text{ is a periodic orbit in } P\}$$

and let

$$\tilde{\mathcal{S}}_\epsilon(P) := \mathcal{S}_\epsilon(P) \cap \Sigma.$$

Every periodic orbit $p \in P$ corresponds to a fixed point $\tilde{p} \in \Sigma$. Observe that

$$\tilde{\mathcal{S}}_\epsilon(P) \subset \bigcup_{\tilde{p} \in W^c(\tilde{q})} W^{cs}(\tilde{p}).$$

Since $W^{cs}(\tilde{p})$ is a manifold of a codimesion at least 2 for all $\tilde{p}$ and $W^c(\tilde{q})$ is a 1-dimensional manifold the measure of $\tilde{\mathcal{S}}_\epsilon(P)$ in $N_\epsilon(P) \cap \Sigma$ is zero. It follows that the measure of $\mathcal{S}_\epsilon(P)$ in $N_\epsilon(P)$ is zero.

Therefore the stable set of P

$$\mathcal{S}(P) := \{x(0) \in B,\ x(t) \to q \text{ as } t \to \infty \text{ and } q \text{ is a periodic orbit in } P\}$$

has measure zero in B being a subset of a countable union of sets of measure zero:

$$\mathcal{S}(P) \subset \bigcup_{n=0}^{\infty} \varphi_{-n}(\mathcal{S}_\epsilon(P)) \cap B.$$

Since the origin is hyperbolic, it is an isolated invariant set and a similar argument as above implies that the stable set of the origin has measure zero in B.

Finally, we note that there at most countably many isolated sets P in the bounded set $U \cap B$. Hence

$$Q = \bigcup_{P \subset (U \cap B)} \mathcal{S}(P)$$

has measure zero. Since B is arbitrary bounded set almost all trajectories converge to a stable, large periodic orbit q in $S \subset M(0)$.

We remark, that in general the center-stable manifold $W^{cs}(\tilde{p})$ is not unique. However, the part of the center-stable manifold which interests us, which is the set of all $x \in W^{cs}(\tilde{p})$ which converge in forward time to $\tilde{p}$, is unique (see [1] for the discussion of center-stable manifolds for flows and [15] Theorem III.7 for maps).

By Proposition 4.8 every large periodic orbit q in $S \subset M(0)$ is a subset of the interior of the set of orthants $\mathcal{O}(V_j)$. There is $2n$ such orthants and since the orbit is large it must go through all $2n$ orthants before it can close up. Therefore this orbit must go from orthant to orthant through the common faces. We used X_s to denote the common face between $\mathcal{O}(V_{s-1})$ and $\mathcal{O}(V_s)$. Recall that X_s is a part of hyperplane $x_s = 0$.

The only thing left to show is that the orbit goes through the orthants in the order of a periodic pattern $\{V_l\}_{l=1}^{2n}$. On the face X_s we have $x_s = 0$ and so from (9)

$$\dot{x}_s = w^*_{s,s-1} g_{s-1}(x_{s-1}).$$

We determine the sign of $\dot{x}_s$. Note that $sign\ x_{s-1} = v^{s-1}_{s-1}$ and the limiting weight $w^*_{s,s-1} = v^s_j v^{s-1}_{j-1}$ for any j and so in particular for $j = s$. Therefore

$$sign\ \dot{x}_s = v^s_s (v^{s-1}_{s-1})^2 \frac{dg_{s-1}}{dx_{s-1}}.$$

Since $v_s^s = sign\, x_s$ in $\mathcal{O}(V_s)$ and $\frac{dg_{s-1}}{dx_{s-1}} > 0$ we get that $sign\, \dot{x}_s = sign\, x_s$. This shows that the only admissible transition is $\mathcal{O}(V_{s-1}) \to \mathcal{O}(V_s)$. This proves the Theorem 1.2. $\square$

5 Cascade of nets

We identify the first equation in (16) as net $\mathcal{M}_0$ and the second equation as net $\mathcal{M}_1$. This notation is in agreement with notation used in [7]. These two nets form a cascade of nets $\mathcal{M}_0 \to \mathcal{M}_1$.

M. Hirsch [7] considered a cascade of nets where each net has a convergent dynamics. His results are not directly applicable since the net $\mathcal{M}_1$ does not have a convergent dynamics. H. Smith [16] considered a case when the net $\mathcal{M}_1$ has convergent dynamics (and has a special form) while the dynamics of $\mathcal{M}_0$ can be convergent or oscillatory. His results are also not directly applicable. In order to use a result of K.Mischaikow, H.Smith and H. Thieme ([12]) below, we need a definition.

Definition 5.1 An (ϵ, τ)-chain from x to y under the flow φ in $\mathbf{R}^n$ is a finite sequence $\{(x_i, t_i)\}_{i=1}^k$ such that

$$\{(x_i, t_i) \in \mathbf{R}^n \times [0, \infty) \mid x = x_1, |\varphi(x_i, t_i) - x_{i+1}| < \epsilon \text{ and } |\varphi(x_k, t_k) - y| < \epsilon\}.$$

If there is a chain from x to y for every ϵ we write $x \geq y$.

A chain recurrent set under the flow φ is defined by

$$\mathcal{R} := \{x \in \mathbf{R}^n \mid x \geq x\}.$$

The union of an equilibrium and a homoclinic orbit to that equilibrium is an example of a chain recurrent set.

Theorem 5.2 (Theorem 1.8, [12]) *Let Φ denotes the flow given by cascade (16) and let φ denotes the flow of (4). Assume that the forward trajectory $\Phi^+(t, x)$ is bounded. Then $\omega = \omega_\Phi(x)$ has the following properties:*

1. *ω is non-empty, compact and connected*
2. *ω is invariant under the flow φ*
3. *ω is chain recurrent for φ*

We now prove the main result.

Proof of Theorem 3.7. Let Λ be the ω-limit set of trajectory through $(\mathbf{x}, W(0))$.

By Theorem 5.2 set Λ is connected, chain recurrent set of (4). Thus Λ has the form (w^*, q) where q is a chain recurrent set of (4). Since the origin is the only equilibrium of the system (4), by Theorem 4.1 the only chain recurrent sets are periodic orbits, the origin or a possible homoclinic to the origin. Since we assume that the origin is hyperbolic the existence of the homoclinic orbit can be ruled out using discrete Lyapunov function argument (see [13] and Appendix I).

Recall that the equilibrium W^* is asymptotically stable in the net $\mathcal{M}_0$. The dimension of the center-stable manifold of Λ in (16) is $r+2n^3$ where r is the dimension of the center-stable manifold of q in system (4) and $2n^3$ is the dimension of the stable manifold of the equilibrium W^*. Thus if the invariant set $q \in U$ then any center-stable manifold of q has codimension at least two in the combined system (16).

Recall, that system (1) is dissipative. It is easy to see that the system (13) is also dissipative. Therefore the combined system (16) is dissipative and there is a bounded set B, which contains the invariant set. Fix such a bounded set $B \subset \mathbf{R}^{2n^3} \times \mathbf{R}^n$. Arguing as in the proof of Theorem 1.2, the set of points

$$Q := \{(y, x) \in B \cap W^{cs}(w^*, q), (y, x) \to (w^*, q) \text{ as } t \to \infty \mid q \in U\}$$

has measure 0. It follows that for almost all $\mathbf{x} \in \mathbf{R}^n$, the trajectory in the system (16) with initial data $(\mathbf{x}, W(0))$ converges to a stable, large periodic orbit exhibiting the periodic pattern $\{V_l\}_{l=1}^{2n}$. $\square$

References

[1] Chow, S-N. & Hale, J. (1982). **Methods of bifurcation theory**, Springer-Verlag, New York, Berlin, Heidelberg.

[2] Gedeon, T. Cyclic feedback systems, accepted to Memoirs of AMS.

[3] Grossberg, S.(1973). Contour enhancement, short term memory, and constancies in reverberating neural networks, *Studies in Applued Math.*,**52**, 217-257.

[4] Grossberg, S.(1978). Competition, decision, and consensus, *Journal of Math. Analysis. and Applications.*, **66**, 470-493.

[5] Hastings, S., Tyson, J. & Webster, D.(1977) Existence of periodic solutions for negative feedback cellular control systems, *J. Differential Equations* **25**, 39-64.

[6] Hebb, D. (1949). **The organization of behaviour**, Wiley, New York.

[7] Hirsch, M. (1989). Convergent activation dynamics in continuous time networks, *Neural Networks*, Vol.2, pp. 331-349.

[8] Hopfield, J.(1982). Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci.* vol.79, pp.2554-2558.

[9] Hopfield, J.(1984). Neurons with graded response have collective computational properties like those of two-state neurons, *Proc. Natl. Acad. Sci.* vol.81, pp.3088-3092.

[10] Gedeon, T. & Mischaikow, K.(1995). Structure of the global attractor for cyclic feedback systems, *J. Dynamics & Differential Equations*, **7**, 141-190.

[11] A. F. Ivanov, J. Losson, Stable rapidly oscillating solutions in delay equations with negative feedback, preprint.

[12] Mischaikow K., Smith H. & Thieme H. (1995). Asymptotically autonomous semiflows: Chain recurrence and Liapunov functions, preprint.

[13] Mallet-Paret, J. & Smith, H.(1990). The Poincarè-Bendixon theorem for monotone feedback systems, *J. Dynamics & Differential Equations* **2**, 367-421.

[14] Mallet-Paret, J. (1996). Stability and oscilation in monotone cyclic systems of differential equations, preprint LCDS #94-17.

[15] Shub, M. (1987).*Global stability of dynamical systems*, Springer-Verlag, New York, Berlin, Heidelberg, Berlin, Paris, Tokyo.

[16] Smith, H.(1991). Convergent and oscillatory activation dynamics for cascades of neural nets with nearest neighbor competitive or cooperative interactions, *Neural Networks*, Vol.4, pp. 41-46.

[17] Troyer, T.(1992). *A Liapunov method for correlational learning in two-layer neural networks.* Doctoral thesis, University of California at Berkeley, Mathematics Department.